

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 15, 2005, 10:06:35 ; Search time 147 Seconds
(without alignments)
4241.070 Million cell updates/sec

Title: AF334735
Perfect score: 1641
Sequence: 1 TCGCCCTTCTCGCGCGCG.....AAAAAAAAAAAAAAAAAAAA 954

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1376875 seqs, 326749119 residues

Total number of hits satisfying chosen parameters: 2753750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DRV=xlp
-Q=/cgn2_1/USPTO_spool_p/AF334735/runat_15022005_094945_10184/app_query.fasta.1.1095
-DB=Published Applications AA -QFWI=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=50 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USPR=AF334735 @CGN 1.1 199 @runat_15022005_094945_10184
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCT05_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description
1	121.5	7.4	697	15	US-10-181-638A-2	Sequence 2, Appli
2	118.5	7.2	382	16	US-10-408-765A-1431	Sequence 1431, Ap
3	114.5	7.0	103	10	US-09-764-891-3819	Sequence 3819, Ap
4	110.5	6.6	298	14	US-10-017-161-2194	Sequence 2194, Ap
5	110.5	6.6	298	15	US-10-292-798-1840	Sequence 1840, Ap
6	108.5	6.6	511	16	US-10-437-963-148275	Sequence 148275,
7	108	6.6	403	14	US-10-092-947A-38	Sequence 38, Appl
8	107.5	6.6	805	15	US-10-369-493-152	Sequence 152, App
9	104	6.3	580	15	US-10-424-599-237561	Sequence 237561,
10	103.5	6.3	532	15	US-10-114-270-178	Sequence 178, App
11	103.5	6.3	535	15	US-10-425-114-64123	Sequence 64123, A
12	103	6.3	197	13	US-10-101-487-511	Sequence 51, Appl
13	103	6.3	197	13	US-10-101-487-114	Sequence 114, App
14	103	6.3	238	14	US-10-163-668-264	Sequence 264, App
15	103	6.3	252	9	US-09-778-927A-42	Sequence 42, Appl
16	102.5	6.2	582	16	US-10-437-963-145629	Sequence 145629,
17	102	6.2	267	9	US-09-778-927A-43	Sequence 43, Appl
18	102	6.2	376	14	US-10-156-761-9889	Sequence 9889, Ap
19	101	6.2	390	15	US-10-108-260A-4549	Sequence 4549, Ap
20	101	6.2	647	9	US-09-815-242-11218	Sequence 11218, A
21	101	6.2	647	15	US-10-282-122A-58476	Sequence 58476, A
22	101	6.2	649	15	US-10-467-248-3	Sequence 3, Appli
23	101	6.2	3067	10	US-09-949-029-18	Sequence 18, Appl
24	100.5	6.1	1487	16	US-10-437-963-161300	Sequence 161300,
25	100.5	6.1	2468	16	US-10-755-889-615	Sequence 615, App
26	99.5	6.1	743	14	US-10-087-454-53	Sequence 53, Appl
27	99	6.0	498	15	US-10-282-122A-72013	Sequence 72013, A
28	99	6.0	521	15	US-10-425-114-70443	Sequence 70443, A
29	99	6.0	670	15	US-10-363-616-339	Sequence 339, App
30	99	6.0	775	10	US-09-934-455-450	Sequence 450, App
31	99	6.0	1537	16	US-10-437-963-183719	Sequence 183719,
32	98.5	6.0	426	13	US-10-087-192-1197	Sequence 1197, Ap
33	98	6.0	273	15	US-10-424-599-166145	Sequence 166145,
34	98	6.0	478	15	US-10-434-599-208282	Sequence 208282,
35	98	6.0	480	9	US-09-347-331-9	Sequence 9, Appli
36	98	6.0	916	16	US-10-408-765A-1222	Sequence 1222, Ap
37	97.5	5.9	197	15	US-10-424-599-174175	Sequence 174175,
38	97.5	5.9	574	16	US-10-437-963-202437	Sequence 202437,
39	97.5	5.9	620	16	US-10-437-963-158544	Sequence 158544,
40	97.5	5.9	429	15	US-10-092-900A-20	Sequence 20, Appl
41	97	5.8	484	15	US-10-369-493-6247	Sequence 6247, Ap
42	97	5.9	995	15	US-10-094-749-2626	Sequence 2626, Ap
43	97	5.9	1261	13	US-10-147-268-2	Sequence 2, Appli
44	97	5.9	1261	14	US-10-338-279-2	Sequence 2, Appli
45	96.5	5.7	438	14	US-10-017-161-2198	Sequence 2198, Ap

ALIGNMENTS

RESULT 1
US-10-181-638A-2
; Sequence 2, Application US/10181638A
; Publication No. US20030207393A1
; GENERAL INFORMATION:
; APPLICANT: Naaby-Hansen, Soren
; APPLICANT: Wolkowicz, Michael
; APPLICANT: Mandal, Arabinda
; APPLICANT: Buer, Sen
; APPLICANT: Herr, John C
; TITLE OF INVENTION: CBP86, A Sperm Specific Protein
; FILE REFERENCE: 00492-07
; CURRENT APPLICATION NUMBER: US/10181,638A
; PRIOR FILING DATE: 2002-10-15
; CURRENT FILING DATE: 2000-01-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-181-638A-2

Alignment Scores:
Pred. No.: 0.00339 Length: 697
Score: 121.50 Matches: 35
Percent Similarity: 46.03% Conservative: 23
Best Local Similarity: 27.78% Mismatches: 43
Query Match: 7.40% Indels: 25
DB: 15 Gaps: 3

AF334735 (1-954) x US-10-181-638A-2 (1-697)

QY 169 ATTCCACAGAGATTTGGGAATCTTCTGAAGGGCTGCACACGGAGATCTTCAGAGAGCAA 228
DB 10 ValProTyrGlyLeuLysThrLeuLeuGluGlyIleSerArgAlaValLeuLysThrAsn 29
QY 229 CCGGACATATACCAAGCTTTTCAGAGCTTATTTTGAGAGCTTCTTCAGAGAAAGAGAG 288
DB 30 ProSerAsnIleAsnGlnPheAlaAlaIlePheGlnGluLeuThrMetTyrArgGly 49
QY 289 AAAACCACTTGTATCCAGCAGATGGGGAGTAAAGTAGACGGCTTCTATACCAAT 348
DB 50 AsnThrThrMetAsp-----IleLysAspLeuValLysGlnPhe 62
QY 349 CATGCATTCGAGGAGCAAGAACCCATCGAGAAAGTGTCTTAAACAAGAA----- 399
DB 63 HisGlnIleLysValGluLysTrpSerGluGlyThrThrProGlnLysLeuGluCys 82
QY 400 -----GAGTCTCAGATA-----TCT 414
DB 83 LeuLysGluProGlyLysThrSerValGluSerLysValProThrGlnMetGluLysSer 102
QY 415 GGGAGGAGGAGAGACATCATGTCACCATCTAGACTCTTCGAGGAGATAGGAAAAA 474
DB 103 ThrAspThrAspGluAspAsnValThrArgThrGluTyrSerAspLysThrThrGlnPhe 122
QY 475 GAAGAGGTGCTGCTGTC 492
DB 123 ProSerValTyrAlaVal 128

RESULT 2

US-10-408-765A-1431
; Sequence 1431, Application US/10408765A
; Publication No. US20040101874A1

GENERAL INFORMATION:

; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1431
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1431

Alignment Scores:
Pred. No.: 0.00916 Length: 382
Score: 118.50 Matches: 38
Percent Similarity: 41.25% Conservative: 21
Best Local Similarity: 26.57% Mismatches: 53
Query Match: 7.22% Indels: 31
DB: 16 Gaps: 4

AF334735 (1-954) x US-10-408-765A-1431 (1-382)

QY 157 ACCCACTACCGAATTCACACAGGATTTGGGAATCTTCTGAAGGGCTGCACACGGAGATT 216
DB 2 SerHisIleGlnIleProProGlyLeuThrGluLeuLeuGlnGlyTyrThrValGluVal 21
QY 217 CTGAGAGAGCAACGGGACATATACCAAGCTTTTCAGAGCTTATTTTGAGAGCTTCTCTA 276
DB 22 LeuArgGlnGlnProProAspLeuValGluPheAlaValGluTyrPheThrArgLeuArg 41
QY 277 GAGAAAAGAGAGAAAACCACTTTGTATCCACAGATGGGGAGTAAAGTAGAGACCGC 336
DB 42 GluAlaArgAlaProAlaSerValLeuProAla-----AlaThrProArg 56
QY 337 TTCTATAACAATCATGATTCGAGGAGCAAGAACCCATCGAGAAAGTGTCTTAAACAA 396
DB 57 GlnSerLeuGlyHisProProGlnProGlyProAspArgValAlaAspAlaLysGly 76
QY 397 GAAGAGTCTCAGATATCTGGGAGGAGGAGAGACATCAGTCACCATCTTAGACTCTTCT 456
DB 77 Asp-----SerGluSer 80
QY 457 GAGGAAGATAGGAAAAGAGAGGTTCTCTGTCAAAATCCAGAGTGCCTTCGGGGA 516
DB 81 GluGluAspGluAspLeuGlu-----ValProValProSerArgPheAsnArg 96
QY 517 CACATAGCCAGAGGAGGCAAGAAATGAAACAATAGTCTTCAAAATGAGGAAAAA 576
DB 97 ArgValSer-----ValCysAlaGluThrTyrAsnProAspGluGlu 110
QY 577 GAGGAAAAAC 585
DB 111 GluGluAsp 113

RESULT 3

US-09-764-891-3819
; Sequence 3819, Application US/09764891
; Publication No. US20030077808A1

GENERAL INFORMATION:

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3819
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-891-3819

Alignment Scores:
Pred. No.: 0.0162 Length: 103
Score: 114.50 Matches: 28
Percent Similarity: 50.00% Conservative: 7
Best Local Similarity: 40.00% Mismatches: 30
Query Match: 6.98% Indels: 5
DB: 10 Gaps: 1

AF334735 (1-954) x US-09-764-891-3819 (1-103)

QY 169 ATTCCACAGAGATTTGGGAATCTTCTGAAGGGCTGCACACGGAGATCTTCAGAGAGCAA 228
DB 30 ValProTyrGlyLeuLysThrLeuLeuGluGlyIleSerArgAlaValLeuLysThrAsn 49
QY 229 CCGGACATATACCAAGCTTTTCAGAGCTTATTTTGAGAGCTTCTTCAGAGAAAGAGAG 288
DB 50 ProSerAsnIleAsnGlnPheAlaAlaIlePheGlnGluLeuThrMetTyrArgGly 69
QY 289 AAAACCACTTGTATCCAGCAGATGGGGAGT-----AAGTAGAGAC 333
DB: 16 Gaps: 4

D	b		201	SerCysPheLeu-----ValPheLeuLeuSerCysLeuLeuAlaPheValLeu	216
Q	y		297	GTTGGTTTTCTCTCTTCTAGAAAGGCTCAAAAATAGCGTGCTGCACAAAAGCTGGTAT	238
D	b		217	SerCysPheLeuAlaPhePhe-----LeuLeuPheLeuSerCysLeuLeuSerCysPhe	234
Q	y		237	ATTGTCCGGTTGTCCTCTCAGAATCTCGCG-----TGTCAGCCCTTCAAGAAG	190
D	b		235	LeuValPheLeuLeuSerCysLeuLeuAlaPheValLeuSerCysPheLeuPhe-----	252
Q	y		189	ATTCCCCAATCCCTTGTCGAATTCGGTAGTGGGTGGAGAAATGGAATCGACATCTTCTT	130
D	b		253	-----PheLeuSerPheLeu-PheLeuSerPhePheLeuSerCysPheLe	267
Q	y		129	GGTAAGAACCTGCTATGGAACCTCTCTCTCCGAGCTGGTCCGCCGGTTCGGGTGT	70
D	b		267	uPhePheHisSerPheLeuLeuSerPheLeuSerPhePheLeuProSer-----	283
Q	y		69	TTTTCTAGTTGCTGGGTAAACGGTTTTTCTT 39	
D	b		284	-PheLeuProSerPheLeuSerPhePheLeu 293	
 RESULT 5					
US-10-292-798-1840					
; Sequence 1840, Application US/10292798					
; Publication No. US2003023583A1					
; GENERAL INFORMATION:					
; APPLICANT: SUWA, MAKIKO					
; APPLICANT: ASAI, KIYOSHI					
; APPLICANT: AKIYAMA, YUTAKA					
; APPLICANT: ASURATANI, HIROYUKI					
; TITLE OF INVENTION: GUANOINSE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS					
; FILE REFERENCE: 084335/166					
; CURRENT APPLICATION NUMBER: US/10/292,798					
; CURRENT FILING DATE: 2002-11-13					
; PRIOR APPLICATION NUMBER: 10/017,161					
; PRIOR FILING DATE: 2001-12-18					
; PRIOR APPLICATION NUMBER: JP 2001-246789					
; PRIOR FILING DATE: 2001-06-18					
; NUMBER OF SEQ ID NOS: 2070					
; SOFTWARE: PatentIn Ver. 2.1					
; SEQ ID NO 1840					
; LENGTH: 298					
; TYPE: PRT					
; ORGANISM: Homo sapiens					
US-10-292-798-1840					
 Alignment Scores:					
Pred. No.: 0.0544 Length: 298					
Score: 110.50 Matches: 75					
Percent Similarity: 39.85% Conservative: 33					
Best Local Similarity: 27.68% Mismatches: 91					
Query Match: 6.57% Indels: 72					
DB: 15 Gaps: 13					
 AF334735 (1-954) x US-10-292-798-1840 (1-298)					
Q	y		759	GTTTCAATGGTAGTCTATTATAACAATGCTCATGACAGATTTTCAACAAGTAACGAAT	700
D	b		65	LeuSerLeuAlaCysLeuLeu-----AlaPhe 73	
Q	y		699	GTTATTTTCAACATCAAATCTCTTCTCAGAGAAGAAATGACATTAAATAAGAGGTG	640
D	b		74	ValLeuSerCysPheProValPheLeuLeuSerPhePheLeuSerPheLeuSerCysLeu	93
Q	y		639	ATGGAT-----TTGGATTATTTTTCATGTTTCTCTGAGGAGTAAACACAGTGTCTCACTT	586
D	b		94	LeuSerCysLeuLeuAlaCysPheArgAlaPheLeu-----LeuSerCysPhe	109
Q	y		585	GTTTTCCTCTTTTTCCTCATTTTGAAGACTATTGTTTTCATTCTTTTCCTTCCTCTCT	526
D	b		110	LeuPheLeuSerPhe---PheLeuSerPheTyrCysPhePheLeuAlaCysPheLeuAla	128

QY 525 GGCTATGTGTCGCCGAGGAGCTTGGATTGTGACGAGCAACCTCTCTTTTCTT 466
Db : : : : :
129 CysLeuLeuAlaPheValLeuSerCys---PheProValPhePheLeuSerPhePheLeu 147
QY 465 APTCTTCTCAGAGAGTCTTAAGATGGTGAATGCTCTC-----TTC 424
Db : : : : :
148 LeuPheLeuSer-----CysLeuLeuSerCysLeuLeuAlaPhe 160
QY 423 CTCCTTCCAGATATCTGAGACTCTTCTGTTAGGATCATTCTCAGGTGGTCTTG 364
Db : : : : :
161 ValLeuSerCysPheLeuAlaPhePheLeuSerPheValSerPheLeuLeuAlaPheLeu 180
QY 363 CTC-----CTC 358
Db : : : : :
181 LeuAspCysPheArgAlaPheLeuLeuSerCysPheLeuSerPhePheCysPhePheLeu 200
QY 357 GAATGCATGATTGTATAGAGCGGTCTTCTACTTCTCCCTTCTCTGATCAAA 298
Db : : : : :
201 SerCysPheLeu-----ValPheLeuLeuSerCysLeuLeuAlaPheValLeu 216
QY 297 GTTGGTTTCTCTCTTCTCTAGAGGCTCTCAAAATAGGCTGTGCAAAAGCTGGTAT 238
Db : : : : :
217 SerCysPheLeuAlaPhePhe-----LeuLeuPheLeuSerCysLeuLeuSerCysPhe 234
QY 237 ATTGTCGGTCTCTCTCAGAACTCTCGCG-----TGTGCGCTCTCAAGAAG 190
Db : : : : :
235 LeuValPheLeuLeuSerCysLeuLeuAlaPheValLeuLeuSerCysPheLeuPhe----- 252
QY 189 APTCCCAAAATCTTGTGGAATTCGGTAGTGGGTGTGGAGATGGAATCGACATCTTCT 130
Db : : : : :
253 -----PheLeuSerPheLeu-PheLeuSerPhePheLeuSerCysPheLeu 267
QY 129 GGTAGAACTGCTATGGAACCTCTCTTCTCCGAGCTGGTGGCGCGTTCGGTGT 70
Db : : : : :
267 uPhePheHisPheLeuLeuSerPheLeuSerPheLeuSerPhePheLeuProSer----- 283
QY 69 TTTTCTAGTGTGGTAACTGCTTTTCTT 39
Db : : : : :
284 -PheLeuProSerPheLeuSerPhePheLeu 293

RESULT 6

US-10-437-963-148275
; Sequence 148275, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 148275
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_48722C.1.p3p
US-10-437-963-148275

Alignment Scores:

Pred. No.:	0.1	Length:	511
Score:	108.50	Matches:	63
Percent Similarity:	39.67%	Conservative:	33
Best Local Similarity:	26.03%	Mismatches:	95
Query Match:	6.61%	Indels:	51

DB: 16 Gaps: 10
AF334735 (1-954) x US-10-437-963-148275 (1-511)
QY 3 GCCCTTCTCGGCGCGCTAGTTTTTTTTTTTAAAGAAAAACGGTTATCCAGCAAC 62
Db : : : : :
31 SerProSerArgSerProSerProArgAsnArgArgAspArgSerProSerPro 50
QY 63 TAGAAAAAACAC-----CGAACCAGCGCCACCGCTCGGA 98
Db : : : : :
51 TyrArgSerArgAspArgSerProSerProTyrArgAspArgArgGlnTrpSer 70
QY 99 -----GAGAAAGGAGGTTCATAGCGAGTCTTACCAAGAGATGTCGATTCC 146
Db : : : : :
71 ProTyrHisArgAspArgGlyArgAspValGluArgGlnTrpAlaArgAspAsp 89
QY 147 ATTCTCCAAACACCCACTACCGAAT-----TCCACAAGATTGGGAA 188
Db : : : : :
90 -----HisProAlaProArgArgGlyGlyAlaGlyAlaGlyAlaTrpSer 106
QY 189 TCTTCTTGAAGGCTGACACCGGAGATCTTGAGAGCAACCGGACATATACAGCTTT 248
Db : : : : :
107 AlaSer-----AspAspAspAspGluGlnLeuLysGly----- 118
QY 249 TGCAGCAGCTTATTTGAGAGCTTCTAGAGAAAGAGAGAAACC----- 294
Db : : : : :
119 ---LeuThr-TyrPheGluTyrArgArgValLysArgGluLysLeuArgLysSerMetLy 137
QY 295 -----AACTTTGATCCAGCAGATGGGGAGTAAAGTAGAAGACCGCTTCTA 341
Db : : : : :
137 sArgCysAlaTrpAsnIleThrProSerProArgArgGluGlyGluAspGluAspTy 157
QY 342 TAACAATCATGCTTCGAGGAGCAAGAACCCACTGAGAAAAGTATCTCTAAACAAGA 401
Db : : : : :
157 rGlyTyrSerAspGluGluGluGluGluGluLysLysLysLysValAl 177
QY 402 GTCTCAGATATCTGGGAGGAGGAGAGACATCAGTCACCATCTTAGACTCTCTGAGGA 461
Db : : : : :
177 aserSerAspLysSerGluGluAspSerLys-----GlySerSerGlucose 193
QY 462 AGATAAGGAAAAAGAGAGTGTCTGCTGTCAAAATCCAAAGCTCCCTCCGGGACACAT 521
Db : : : : :
193 rAspSerGlyGluSerAsp-SerLeuSerAspSerLys-----SerAspAspThrA 211
QY 522 AGCCAGAGAGGAGCAAGAAAATGAAACAAATAGTCTTCAAAATG-----AGGA 572
Db : : : : :
211 rArgLysLysLysGlyArgLysGlySerHisArgSerSerLysArgSerArgHisArgA 231
QY 573 AAAAGAGGAAAAAAGTGAGGACACTGGTTTTTACCTCCAGGAAACATGAAAAATAATCCA 632
Db : : : : :
231 rArgHisHisSerSerAspThrGluGlyAspAspAsnSerLysAlaGluAspSerG 251
QY 633 AA 634
Db : :
251 lu 251

RESULT 7

US-10-092-947A-38
; Sequence 38, Application US/10092947A
; Publication No. US20030134353A1
; GENERAL INFORMATION:
; APPLICANT: WOLFF, Anne M
; APPLICANT: APPEL, Karen F
; APPLICANT: PETERSEN, Jesper F
; APPLICANT: POULSEN, Ulla
; APPLICANT: ARNAU, Jose
; APPLICANT: JACOBSEN, Mette D
; TITLE OF INVENTION: MUCOR RECOMBINANT GENE EXPRESSION
; FILE REFERENCE: WOLFF-3
; CURRENT APPLICATION NUMBER: US/10/092,947A
; CURRENT FILING DATE: 2002-12-27
; PRIOR APPLICATION NUMBER: US 60/274,650
; PRIOR FILING DATE: 2001-03-12

Percent Similarity: 36.89% Conservative: 24
Best Local Similarity: 25.24% Mismatches: 72
Query Match: 6.34% Indels: 58
DB: 15 Gaps: 9

AF334735 (1-954) x US-10-424-599-237561 (1-580)

```
QY 64 AGAAAAACACCGGACCG-----GGCGGCACAGCTCGGAGAGA 102
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 205 ArgLysAspGluGlyThrValArgLeuValAlaLysGlyGlyThrArgSerAspGly 224
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 103 AAGGAGGTTCCATAGCAGCTTCTACCAAGAAGATGTCATCCATCTTCCAAACCCAC 162
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 225 AsnGlu-----LysLeuLysAspMetGlyLeuAspValLysLysLeuAsp 239
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 163 TACCGAATCCACAGGATTTGGGAATCTTCTGAAGGCTGACAGCGAGATTCGTAGA 222
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 240 -----GlyArgGlyLeuGlnAspGluValArgProIleGlyAsn 252
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 223 GAGCAACCG-----GACAATATACAGCT 246
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 253 AlaAlaProValGlnAsnHisValGlyAsnPheHisProArgValAspGlyIleProLys 272
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 247 TTTGCGACGACCTATTTTTCAGAGCCTTCTAGAGAAAGAGAGAAACCAACTTTGATCCA 306
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 273 LeuLeuGlyLysTyPheGluArgAsnLeuGluAla-----Thr 285
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 307 GCAGATGGGGAGTAGTAGAGACCGCTTCTATAACATCATGCTTCGAGGAGCAA 366
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 286 ValGluGlyLysGluLysValLys-----GluLysLys 296
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 367 GAACACCTGAGAAAGTATCTTAACAAGAGAGTCTCAGATATCTGGGAGGAGAA 426
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 297 AspGluLysGluLysAlaLysGluLysThrGluGluArgLysValLysGluLysAsp 316
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 427 GAGACATCATGTCACCATCTTACTGAGTCTTCTGAGGAA-----GAT 465
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 317 GluGlyLysGluLysValLysGluLysLysAspAspLysArgAspLysArgLysAsp 336
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 466 AGGAAAAAGAGAGTGTCTGCTCAAAATCCAGTGCCTCCGGGAGACATAGCC 525
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 337 LysGluLysGluLysLysGlyHisGlyLysAspLysAspArgAspLysLysLys 356
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 526 AGAGAGGAGCGAAG-----AAAATGAACAAATAGTCTTCAAAATCAGGAGAAA 576
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 357 GluGluLysAlaLysGluHisGlyGluLeuLysThrGlu-----GlnAsnLysLeuLys 375
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 577 GAGGAAAAACAGTGAGGA 594
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 376 GluSerAsnLysValGly 381
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 10

```
US-10-114-270-178
; Sequence 178, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Patturajan, Meera
; APPLICANT: Liu, Ziaohong
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gorman, Linda
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Smithson, Glenna
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gerlach, Valerie

Alignment Scores:
Pred. No.: 0.321 Length: 532
Score: 103.50 Matches: 41
Percent Similarity: 42.86% Conservative: 25
Best Local Similarity: 26.62% Mismatches: 59
Query Match: 6.31% Indels: 29
DB: 15 Gaps: 7

AF334735 (1-954) x US-10-114-270-178 (1-532)
QY 181 TTTGGGATCTTCTTGAGGGCTGACGCGGAGATTCTGAGAGACCAACCGACAATATA 240
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 164 PheGly-----GluSerLeuLeuSerAspMetLeuAspPheProAspThrLeu 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 CCAGCTTTTGAGCAGCCTATTTTGAG-----AGCCTT 273
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 ProSerProGluAlaLeuLysPheLysLysLeuValLysAsnLysLysLysLysLeu 200
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 274 CTAGAGAAAAGAGAGAAAACCACTTTGATCCA-----GCAGATGGGGAGT 321
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 201 LysGluThrHisGluArgLysGlySerAspLysArgGlyLysValGluGluGluGlu 220
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 322 AAGGTAGAAGACCGCTTCTATAACAATCATCATCGAGGAGCAAGACCACTGAGAAA 381
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 221 GluValAlaAspGly-----GluGluGluGluGluGluGluGlu 233
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 382 AGTGATCCTAAACAAGAGAGAGTCTCAGATATCTCGGAGGAGGAGAGACATCAGTCACC 441
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```
; PRIOR APPLICATION NUMBER: IL 134453
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: IL135341
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)..(252)
; OTHER INFORMATION: Xaa = any amino acid, unknown, or other
US-09-778-927A-42

Alignment Scores:
Pred. No.: 0.294 Length: 252
Score: 103.00 Matches: 23
Percent Similarity: 48.75% Conservative: 10
Best Local Similarity: 36.25% Mismatches: 26
Query Match: 6.28% Indels: 15
DB: 9 Gaps: 1

AF334735 (1-954) x US-09-778-927A-42 (1-252)
QY 351 TGCATTTCGAGGAGCAAGAACCTGAGAAAAGTGTCTTCAACCAAGAGAGTCTCAGAT 410
Db CysMetArgArgThrLysGlnVal-GluLysAsnAspAspGlnLysIleGluGlnAs 23
QY 411 ATCTGGGAGGAGGAGAGACATCAGTCACCATCTTAGACTCTCTTCGAGGAAGATAAGA 470
Db 23 pGlyLeuLysProGlu-----As 29
QY 471 AAAAGAGAGGTTGCTGCTGTCCTCAAAATCCAAAGTGCCTTCGGGGACACATAGCCAGAGA 530
Db 29 physalaHisLysAlaalaThrLysIleGlnAlaSerPheArgGlyHisIleThrArgLys 49
QY 531 GGAGGCAAGAAAATGAAACAAATAGTCTTCAAAATGAGGAAAAGAGGAAAACAAAG 588
Db 49 slyLeuLysGlyGluLysLysAspValGlnAlaAlaGluAlaGluAlaAsnLys 68

RESULT 16
US-10-437-963-145629
; Sequence 145629, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 145629
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; LOCATION: (1)..(582)
; OTHER INFORMATION: Xaa = any amino acid, unknown, or other
US-10-437-963-145629

Alignment Scores:
Pred. No.: 0.415 Length: 582
Score: 102.50 Matches: 35
Percent Similarity: 45.08% Conservative: 20
```

```
Best Local Similarity: 28.69% Mismatches: 38
Query Match: 6.25% Indels: 29
DB: 16 Gaps: 5

AF334735 (1-954) x US-10-437-963-145629 (1-582)
QY 253 GCAGCCTATTTCGAGAGCCTTCTA-----GAGAAAGAGAGAAACCAACTTTCATCCA 306
Db 6 AlalystrpIleLysSerValLeuGlyLysLysSerAlaLysSerAsn----- 22
QY 307 GCAGAATGGGGAGTAAAGTAGAGAC-----CGCTTCTATAACAATCATGCTCATTCGAG 360
Db 23 -----SerThrLysAlaLysAspLeuAlaLysAlaLysAsnLysProValLeu 39
QY 361 GAGCAAGAACCACTGAGAAAAGTGTCTTCAACCAAGAGAGTCTCAGATATCTCGGAAG 420
Db 40 SerGluAspProValIleSerGluProAlaLeuValAsnSerHisAsnAspGlyAsn 59
QY 421 GAGGAGAGAGATCAGTC----- 438
Db 60 AlaGluAsnCysLysLeuProAsnGlyValAlaValGluAlaMetGlyGlnGlyValGlu 79
QY 439 -----ACCATCTTAGACTCT-----TCTGAGGAAGATAAGGAAAAAGAGAGGTT 483
Db 80 AsnGlnAsnIleValGlySerLysAlaProThrSerProGluLysLeuSerGluGluLeu 99
QY 484 GCTGCTGTCAAAATCCAAAGTGCCTTCGGGGACACATAGCCAGAGAGGCAAGAA 543
Db 100 AlaAlaValLysAlaGlnAlaAlaPheArgGlyTyrlLeuAlaAtgArgAlaPheArgAla 119
QY 544 ATGAAA 549
Db 120 LeuLys 121

RESULT 17
US-09-778-927A-43
; Sequence 43, Application US/09778927A
; Patent No. US20020068342A1
; GENERAL INFORMATION:
; APPLICANT: KHOSRAVI, Rami et al.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID AND AMINO ACID SEQUENCES AND NOVEL
; TITLE OF INVENTION: VARIANTS OF ALTERNATIVE SPLICING
; FILE REFERENCE: 2786-0160P
; CURRENT APPLICATION NUMBER: US/09/778,927A
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: IL 134453
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: IL135341
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(267)
; OTHER INFORMATION: Xaa = any amino acid, unknown, or other
US-09-778-927A-43

Alignment Scores:
Pred. No.: 0.377 Length: 267
Score: 102.00 Matches: 27
Percent Similarity: 49.30% Conservative: 8
Best Local Similarity: 38.03% Mismatches: 22
Query Match: 6.22% Indels: 14
DB: 9 Gaps: 1

AF334735 (1-954) x US-09-778-927A-43 (1-267)
QY 376 GAGAAAAGTGTCTTCAAAACAGAGAGGTCTCAGATATCTGGGAGGAGAGAGATCA 435
```

```
Db 41 GluLysAsnAspAspAspGlnLysIleGluInAspGlyIleLysProGlu----- 57
QY 436 GTCCACATCTTAGACTCTTCTGAGGAGATAGAGAAAGAGAGAGTTGCTGCTGCTCAA 495
Db 58 -----AspLysAlaHisLysAlaAlaThrLys 66
QY 496 ATCCAAAGCTGCCTTCGGGGACACATAGCAGAGAGGAGGCAAGAAATGAAAAACAAT 555
Db 67 IleGlnAlaSerPheArgGlyHisIleThrArgLysLysLeuLysGlyGluLysLysAsp 86
QY 556 ACTCTTCAAAATGAGCAAAAGAGAGAAACAAG 588
Db 87 AspValGlnAlaAlaGluAlaGluAlaAsnLys 97

RESULT 19
US-10-156-761-9889
; Sequence 9889, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9889
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9889

Alignment Scores:
Pred. No.: 0.413 Length: 376
Score: 102.00 Matches: 39
Percent Similarity: 39.38% Conservative: 24
Best Local Similarity: 24.38% Mismatches: 55
Query Match: 6.22% Indels: 42
DB: 14 Gaps: 5

AF334735 (1-954) x US-10-156-761-9889 (1-376)
QY 169 ATTCACAAAGATTGGGAATCTTCTGAAGGCTGACACGGAGATTCTGAGAGAGCAA 228
Db 58 MetProGln---PheAlaGlnLeuGlnLeuArgGlyGluValLeuAspAlaGly 76
QY 229 CCGGACAATATACCAAGCTTTGCACAGCCCTATTTTGAGAGCTTCTA----- 276
Db 77 ArgLysAlaValThrAlaAlaAlaAspArgGlyMetSerSerLeuAlaAspAlaLeuSer 96
QY 277 -----GAGAAAGAGAG 288
Db 97 AspArgThrAlaArgLeuGlyLysLysGlyGluGluGluGlyGluGluGluGly 116
QY 289 AAAACCAATTGATCCAGCAAGATGGGGAGTAGAGTAGAGACCGCTTCTATAACAAT 348
Db 117 GluGluGluTyrAlaProAspGluGlyAlaGluGlyGluGlu----- 131
QY 349 CATGCATTCGAGGACCAAGCAACCTCAGAAAAGTGATCTTAACAAGAGAGTCTCAG 408
Db 132 GluGluGluGluGluGluGluProGluGluGluGluProGluGluGluGluGlu 151
QY 409 ATATCTGGGAGGAGAGAGACATCAGTCACCATCTTAGACTCTTCTGAGGAGAGATAAG 468
|||||
|||||
```

```
Db 152 ProGluAlaAlaLysGluGlu-----GluGluGluGluGlu 163
QY 469 GAAAGAGAGAGTTGCTGCTGCTCAAAATCCAAGCTGCTTCCGGGGACACATAGCCAGA 528
Db 164 GluGluGluGluGlnPro-----GluGlyGluTyrGluGlu 175
QY 529 GAGGAGGCAAGAAATGAAACAAATAGTCTTCAAAATGAGGAAAAAGAGGAAACAAG 588
Db 176 GluGluGluGluGluGluGluGluProGluGluGluGluGluGluGluGlu 195

RESULT 19
US-10-108-260A-4549
; Sequence 4549, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4549
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4549

Alignment Scores:
Pred. No.: 0.526 Length: 390
Score: 101.00 Matches: 42
Percent Similarity: 41.32% Conservative: 27
Best Local Similarity: 25.15% Mismatches: 56
Query Match: 6.15% Indels: 42
DB: 15 Gaps: 8

AF334735 (1-954) x US-10-108-260A-4549 (1-390)
QY 181 TTTCGGAAATCTTCTTGAAGGCTGACACGGGCTGATCTGAGAGACACCGACAATATA 240
Db 9 PheGly-----GluSerLeuLeuSerAspMetLeuAspPheProAspThrLeu 25
QY 241 CCAGCTTTTGACGACGCTATTTTGAG-----AGCCTT 273
Db 26 ProSerProGluAlaLeuLysPheLysIleLeuValLysAsnLysLysIleGlyThrLeu 45
QY 274 CTAGAGAAAAGAGAGAAAACCACTTTGATCCA-----GCAGAAATGGGGAGT 321
Db 46 LysGluThrHisGluArgLysGlySerAspLysArgGlyLysValGluGluTyrGluGlu 65
QY 322 AAGTAGAAGACCGCTTCTATACAAATCATGCATTCAGGAGCAAGAACCCCTGAGAAA 381
Db 66 GluValAlaAspGly-----GluGluGluGluGluGluGlu 78
QY 382 AGTCATCTTAAACAAGAGAGTCTCAGATATCTCGGAAGGAGGAGAGACATCAGTCACC 441
Db 79 GluGluGluGluGluGluGluAspLysPhe-----LysGluSerGluValLeuGluSer 96
QY 442 ATCTTAGACTCTTCTGAGGAGATAGAGAA----- 471
Db 97 ValLeu---GlyAspAsnGlnAspLysGluThrGlyValLysLysLeuProGlyValMet 115
QY 472 -----AAAGAGAGTTGCTGCTGCTCAAAATCCAAGCTGCTTCCGGGGACACATA 522
Db 116 LeuPheLysLysLysThrArgLysLysLysIleAlaLeuAlaLeuSerAspLeuVal 135
QY 523 GCCAGAGAGAGGCAAGAAAAATGAAAAACA-----AATAGTCTTCAAAATGAGGAA 573
Db 136 IleTyrThrLysAlaGluLysPheLysSerPheGlnHisSerArgLeuTyrGlnGlnPhe 155
QY 574 AAAGAGGAAAACAGTCAGGA 594
Db 156 AsnGluAsnAsnSerIleGly 162
|||||
|||||
```


QY 277 GAGAAAGAGAGAAACCACTTTGATCCAGCAGAGTGGGGAGTAAGTAGAACCGC 336
Db 527 ---LysGlnGlnGlnAlaAsnPhe-----TpaLaserLys----- 537
QY 337 TTCTATAACAATCATGCTTCAGGAGGAGCAAGAACCCCTGAGAAAGTGATCCTAAACAA 396
Db 538 -----AlaValGluGluGlnAlaLysAlaLysLysSerGluProLeuLys 552
QY 397 GAAGAGTCT-----CAGATATCTGGG 417
Db 553 GluGluSerAlaValLysAsnAspArgThrSerLysProLysSerValLysLeuSerTyr 572
QY 418 AAGGAGGAA-----GAGACATCACTC 438
Db 573 LysGluGlnArgGluLeuGluGlnLeuProGlnLeuLeuGluGluGluThrLysIle 592
QY 439 ACCATCTTAGACTCTCTCAGGAAGAT-----AAGGAAAAAGAGAGGTTGCT 486
Db 593 ThrValLeuGlnAlaGluIleAlaAspProAlaPhePheGlnGlnAlaHisAspIleThr 612
QY 487 GCTGTCAAAATCCAGCTGCTTCGCGGGACACATAGCCAGAGAGGCGCAAGAAATG 546
Db 613 AspAlaLysLeuLysAla-----LeuAlaAspThrGluAla---GluLeu 626
QY 547 AAAACAAATAGTCTTCAAAATCAGCAAAAGAGGAGGAAACAG 588
Db 627 GluThrAlaPheLeuArgTrpGluGluLeuGluGluLysLys 640

RESULT 22

US-10-467-248-3

; Sequence 3, Application US/10467248

; Publication No. US20040086905A1

; GENERAL INFORMATION:

; APPLICANT: DAS, Debopriya; YAO, Monique G.;

; APPLICANT: ARVIZU, Chandra S.; BAUGHN, Mariah R.;

; APPLICANT: LU, Yan; HAPALIA, April J.A.;

; APPLICANT: CHAWLA, Navinder K.; GRIFFIN, Jennifer A.;

; APPLICANT: LU, Dying Aina M.; YUE, Henry;

; APPLICANT: DING, Li; ELLIOTT, Vicki S.;

; APPLICANT: FORTSHE, Ian J.; RAMKUMAR, Jayalaxmi;

; APPLICANT: GANDHI, Aneeta R.; ISON, Craig H.;

; APPLICANT: WARREN, Bridget A.; TANG, Y. Tom;

; APPLICANT: EMERLING, Brooke M.; HONCHELL, Cynthia D.;

; APPLICANT: LYNE, Michael; BARROSO, Ines;

; TITLE OF INVENTION: LIPID-ASSOCIATED MOLECULES

; FILE REFERENCE: PI-0358 USN

; CURRENT APPLICATION NUMBER: US/10/467,248

; PRIOR FILING DATE: 2003-08-06

; PRIOR APPLICATION NUMBER: PCT/US02/03813

; PRIOR FILING DATE: 2002-02-06

; PRIOR APPLICATION NUMBER: US 60/266,910

; PRIOR FILING DATE: 2001-02-06

; PRIOR APPLICATION NUMBER: US 60/276,891

; PRIOR FILING DATE: 2001-03-16

; PRIOR APPLICATION NUMBER: US 60/276,855

; PRIOR FILING DATE: 2001-03-16

; PRIOR APPLICATION NUMBER: US 60/279,760

; PRIOR FILING DATE: 2001-03-28

; PRIOR APPLICATION NUMBER: US 60/283,818

; PRIOR FILING DATE: 2001-04-13

; PRIOR APPLICATION NUMBER: US 60/285,405

; PRIOR FILING DATE: 2001-04-20

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PERL Program

; SEQ ID NO 3

; LENGTH: 649

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No: 72852842CD1

US-10-467-248-3

Alignment Scores: 0.604 Length: 649
Pred. No.: 101.00 Matches: 42
Score: 41.32% Conservative: 27
Percent Similarity: 25.15% Mismatches: 56
Best Local Similarity: 6.15% Indels: 42
Query Match: 15 Gaps: 8
DB: 15

AF334735 (1-954) x US-10-467-248-3 (1-649)

QY 181 TTGGGATCTTCTTGAAGGCTGACACGCGAGATTTCTGAGAGAGCAACCGGACAAATATA 240
Db 268 PheGly-----GluSerLeuLeuSerAspMetLeuAspPheProAspThrLeu 284
QY 241 CCAGCTTTTTCAGAGCGCTATTTTCAG-----AGCCTT 273
Db 285 ProSerProGluAlaLeuLysPheLysIleLeuValLysAsnLysLysIleGlyThrLeu 304
QY 274 CTAGAGAAAGAGAGAAACCACTTTTCATCCA-----GCAGATGGGGAGT 321
Db 305 LysGluThrHisGluArgLysGlySerAspLysArgGlyLysValGluGluTrpGluGlu 324
QY 322 AAGGTAGAGAGCGCTTCTATAACAATCATGCATTCGAGGAGCAAGAACCCACCTGAGAAA 381
Db 325 GluValAlaAspGly-----GluGluGluGluGluGluGlu 337
QY 382 AGTGATCTTAAACAAGAGAGATCTCAGATATCTCGGAGAGGAGAGAGACATCAGTCACC 441
Db 338 GluGluGluGluGluGluGluAspLysPhe-----LysGluSerGluValLeuGluSer 355
QY 442 ATCTTAGACTCTCTGAGGAGAGATAAGGAA----- 471
Db 356 ValLeu---GlyAspAsnGlnAspLysGluThrGlyValLysLysLeuProGlyValMet 374
QY 472 -----AAAGAGAGAGGTTGCTGCTCTCAAAATCCAAAGCTGCTTCGCGGGACACATA 522
Db 375 LeuPheLysLysLysLysThrArgLysLeuLysIleAlaLeuAlaLeuSerAspLeuVal 394
QY 523 GCCAGAGAGGAGGCAAGAAATGAAACA-----AATAGTCTTCAAAATGAGGAA 573
Db 395 IleTyrThrLysAlaGluLysPheLysSerPheGlnHisSerArgLeuTyrGlnGlnPhe 414
QY 574 AAAGAGGAGAAACAAGTCAGGA 594
Db 415 AsnGluAsnAsnSerIleGly 421

RESULT 23

US-09-949-029-18

; Sequence 18, Application US/09949029

; Publication No. US20030134278A1

; GENERAL INFORMATION:

; APPLICANT: Karpen, G.H.

; APPLICANT: Dobie, K.W.

; APPLICANT: Kennedy, C.D.

; APPLICANT: Velasco, V.M.

; APPLICANT: McGrath, T.L.

; APPLICANT: Weko, J.

; APPLICANT: Patterson, R.W.

; TITLE OF INVENTION: Identification of chromosome inheritance modifiers in *Drosophila*

; FILE REFERENCE: 1211.015US1

; CURRENT APPLICATION NUMBER: US/09/949,029

; PRIOR FILING DATE: 2001-09-07

; PRIOR APPLICATION NUMBER: US 60/231,178

; PRIOR FILING DATE: 2000-09-07

; NUMBER OF SEQ ID NOS: 149

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 18

; LENGTH: 3067

; TYPE: PRT

; ORGANISM: *Drosophila melanogaster*

US-09-949-029-18

Db 1386 SerThrSerAlaProSerProGluLys 1394

RESULT 25

US-10-755-889-615

; Sequence 615, Application US/10755889

; Publication No. US20040171823A1

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB

; TITLE OF INVENTION: PATHWAY

; FILE REFERENCE: D0284 NP

; CURRENT APPLICATION NUMBER: US/10/755,889

; CURRENT FILING DATE: 2004-01-13

; PRIOR APPLICATION NUMBER: U.S. 60/440,068

; PRIOR FILING DATE: 2003-01-14

; PRIOR APPLICATION NUMBER: U.S. 60/469,757

; PRIOR FILING DATE: 2003-05-12

; NUMBER OF SEQ ID NOS: 823

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 615

; LENGTH: 2468

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-755-889-615

Alignment Scores:

Pred. No.:	0.976	Length:	2468
Score:	100.50	Matches:	45
Percent Similarity:	41.85%	Conservative:	32
Best Local Similarity:	24.46%	Mismatches:	73
Query Match:	6.12%	Indels:	34
DB:	16	Gaps:	5

AF334735 (1-954) x US-10-755-889-615 (1-2468)

QY 129 CAAGAAGATGTCGATTCCTCCACACACCCACTACCGAATTCACCAAGGATTGGGAA 188

Db 518 GlnLysAspLeuThrGlyGlnValProThrProValValLysGlnThrLysLeuLysGln 537

QY 189 TCTTCTTGAAGGCTGACGACGAGATTCGTAGAGAGCAACCGGACAAATATACCGACTTT 248

Db 538 -----ArgAlaAspSerArgGluSerLeuLysProAlaAlaLysProLeuProSer 554

QY 249 TGCACGACCTATTTTGAGGCTTCTAGAGAAAGAGAGAAACCAACTTTGATCCAGC 308

Db 555 LysSer-ValArgLysGluSerLysGluGluThrProGluValThrLysValAsnHisVa 574

QY 309 AGAATGGGGAGTAGAGTACGAGAC----- 333

Db 574 lGluLysProProLysValGluSerLysGluLysValMetValLysLysAspLysProVa 594

QY 334 -CGCTTCTATACATCATGCAATTCGAGAGAGCAACACCTGAGAAAGTGATCTTAA 392

Db 594 llyserThrGluThrLysProSerValThrGluLysGluValProSerLysGluGluProSe 614

QY 393 ACAAGAAGTCTCAGATATCTGGAGAGAG----- 423

Db 614 rProValLysAlaGluValAlaGluLysGlnAlaThrAspValLysProLysAlaAlaLy 634

QY 424 -----GAAGACATCATGCTACCATCTTAGACTCTTCTGAGGAA---GA 464

Db 634 sGluLysThrValLysLysGluThrLysValLysProGluAspLysLysGluLysGlu 654

QY 465 TAAGGAAAAGAGAGTTGCTGCTGCTCAAAATCCAAAGCTGCTTCCGGGACAC----- 519

Db 654 uLysProLysLysGluValAlaLysGluAspLysThrProLysLysGluGluLy 674

QY 520 -ATACCCAGAGGAGGCAAGAAATGAAACAATAGTCTTCAAAATGAGGAAAGA 578

Db 674 sProLysLysGluGluValLysLysGluValLysLysGluLysLysGluLysLy 694

QY 579 GGAAAAACAAG 588

Db 694 sGluProLys 697

RESULT 26

US-10-087-464-53

; Sequence 53, Application US/10087464

; Publication No. US20030059436A1

; GENERAL INFORMATION:

; APPLICANT: Chishti, Athar

; APPLICANT: Oh, Steven

; APPLICANT: Liu, David

; APPLICANT: Goel, Vikas

; APPLICANT: Li, Xuerong

; TITLE OF INVENTION: Band 3 Antigenic Peptides, Malaria Polypeptides and Uses Thereof

; FILE REFERENCE: S1237/7019

; CURRENT APPLICATION NUMBER: US/10/087,464

; CURRENT FILING DATE: 2002-03-01

; PRIOR APPLICATION NUMBER: US 06/272,930

; PRIOR FILING DATE: 2001-03-02

; NUMBER OF SEQ ID NOS: 59

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 53

; LENGTH: 743

; TYPE: PRT

; ORGANISM: Plasmodium falciparum

US-10-087-464-53

Alignment Scores:

Pred. No.:	0.887	Length:	743
Score:	99.50	Matches:	37
Percent Similarity:	41.06%	Conservative:	25
Best Local Similarity:	24.50%	Mismatches:	42
Query Match:	6.06%	Indels:	47
DB:	14	Gaps:	5

AF334735 (1-954) x US-10-087-464-53 (1-743)

QY 148 TTCTCCAAACCCACTACCGAATTCACCAAGGATT---GGGAATCTTTGAAGGG--- 201

Db 622 TyrThrAsnGlyAsnLysAsnIleGlnGlnIlePheGlnGlnAsnIleLeuGluAsnAsp 641

QY 202 -----CTGCACGCGAGATTCTGAGAGCAACCGGCAATATATACCGCTTTT 249

Db 642 ValLeuAsnGlnGluThrGluGluMetGluLysGlnValGluAlaIleThrLysGln 661

QY 250 GCACGACGCTATTTGAGAGCCTTCTAGAGAAAGAGAGAAACCAACTTTGATCCAGCA 309

Db 662 lIleGluAlaGluValAspAlaLeuAlaProLysAsnLysGlu----- 675

QY 310 GAATGGGGAGTAGAGACCGCTTCTATAACAATCATGCTTCGAGGAGCAAGAA 369

Db 676 -----GluGluGluLys 679

QY 370 CCACCTGAGAAAGTGATCTTAAACAAGAAGAGTCTCAGATATCTGGGAAGGAGAGAG 429

Db 680 GluLysGluLysGluLysGluLysGluLysGluLysGluLysGluLysGlu 699

QY 430 ACATCAGTCACCTCTTAGACTCTTCTGAGGAAGTAAGGAAAGAGAGGTTGCTGCT 489

Db 700 LysGlu-----GluLysGluLysGluLysGluLysGluLysGlu----- 712

QY 490 GTCAAAATCCAAGCTGCTTCCGGGGACACATAGCCAGAGAGAGGCAAGAAATGAA 549

Db 713 -----LysGluGluGluLysGluLys 720

QY 550 ACAAAATAGTCTTCAAAATGAGGAAAAAGAGAA 582

Db 721 GluGluGluGlnGluGluGluGluGluGluGlu 731

RESULT 27

US-10-282-122A-72013

; Sequence 72013, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

124 p-GluLysProAsnGlnSerProAspPheMetalGluTyrTyrAlaArgLysAlaGluL 144

281 AAAGAGAGAAAAACCAACTTTGAT---CCAGCAGAATGGGGAGTAAGGTAGAAGACCGCT 337

144 euAlaGlnLysValGluSerGlnLysProAlaGlu----- 155

338 TCTATAACAATCATGCANTTCGAGAGCAAGCAACCACTCGAAGAAAGTGATCCTTAACAAG 397

156 -----ThrAlaGluGlnGluSerGlnThrLeuSerAspGlnLysGluT 171

398 AAGAGTCTCAGATATCTGGGAGGAGAGACATCAGTCACCATCTTAGACTCTTCTG 457

171 hrGluThrAlaValAsnGlnSerGluGluAlaSerSerGlnIleAlaGluSerValG 191

458 AGGAAGATAAGGAAAAAGAGAGTCTGCTGTCAAAATCCAAAGTCGCCTTC 510

191 luserGluGluGluLysTyrAsnArgSerLeuLysLysThrArgThrGlyPhe 208

RESULT 28

US-10-425-114-70443

Sequence 70443, Application US/10425114

Publication No. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E

APPLICANT: Tabaska, Jack E

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCES: 38-21(53313)B

CURRENT APPLICATION NUMBER: US/10/425,114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 70443

LENGTH: 521

TYPE: PRT

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Clone ID: UC-ZMROB73035P03_FLI.pep

US-10-425-114-70443

Alignment Scores:

Pred. No.: 0.904 Length: 521

Score: 99.00 Matches: 41

Percent Similarity: 42.03% Conservative: 17

Best Local Similarity: 29.71% Mismatches: 53

Query Match: 6.03% Indels: 27

DB: 15 Gaps: 5

AF334735 (1-954) x US-10-425-114-70443 (1-521)

QY 186 GAATCTTTGAGGGGTGACACGGCAGATTCTGAGAGAGCAACCGGACAAATATACCAGC 245

DB 101 GluSerSerAspSerAspThrAspGluAspGlu-GluPro----- 115

QY 246 TTTTGCAGCAGCTATTTTTCAGAGCCCTTCAGAGAAAAGAGAGAAACCAACTTTGTATCC 305

DB 116 -----GlnLysGlnLysGluAlaLeuSerAl 125

QY 306 AGCAGAATGGGGAGTAAAGTAGAAGACCGCTTCTATAACAATATGCATTCGAGAGCA 365

DB 125 aAlaLysLysGluSerSerSerGluAspGluAspAspSerSerGluGluSerSerAspAs 145

QY 366 AGAACCA-----CCTGAGAAAAGTGATCTCTAAACAAGAGAGCTTCAGATATCTGG 416

DB 145 pGluProThrLysValGluGluLysLysAlaProLysValSerGluAsn-----SerGl 163

QY 417 GAAGGAGGAGAGACATCAGTCACCATCTTAGACTCTTCTGAGGAAGATAGGAAAAAGA 476

DB 163 ySerGluAspGluSerSer-----GluAspGluSerAspLysAspSerGl 178

```
QY 477 AGAGTTCCTGCTGTCACAAATCCAAAGTCCCTCCGGGACACATAGCCAGAGAGGAGGC 536
    |||||
Db 178 uGluProAlaenThrProLysLysAlaAlaValHisAlaSerGluLysLysThrAlaTh 198
    :|||
QY 537 AAAGAAATGAAACA---AATAGTCTTCAAAATGAGGAAAGAGGAAAC 585
    |||||
Db 198 rlysgluProLysThrProThrGlySerGlnSerGluSerThrGluValAsn 215

RESULT 29
US-10-363-616-339
; Sequence 339, Application US/10363616
; Publication No. US20040044181A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-113 (793)
; CURRENT APPLICATION NUMBER: US/10/363,616
; PRIOR FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 09/654,935
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 490
; SEQ ID NO 339
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-363-616-339

Alignment Scores:
Pred. No.: 0.968 Length: 670
Score: 99.00 Matches: 71
Percent Similarity: 39.35% Conservative: 62
Best Local Similarity: 21.01% Mismatches: 120
Query Match: 6.03% Indels: 85
DB: 15 Gaps: 13

AF334735 (1-954) x US-10-363-616-339 (1-670)
QY 66 AAAAACAACCGAACCGCGGCACAGCTCGAGAGAAAGGAGTTCCATA----- 116
    |||||
Db 70 LysAsnLysLysGlyLysAlaArgLysLysGlyLysLysGlyLysLysGlyLysLysGln 89
    :|||
QY 117 -----GGCAGTCTTTACCAAGAAGATGTGATTCATTCTCCAAACACCCACTA 164
    |||||
Db 90 AlaAsnLysThrAlaSerSerGlySerSerAspLysAspSerSerAlaGluSerSerAla 109
    :|||
QY 165 CGAATTCACAGAGATTGGGAATCTCTTGAAGCGGTGACACCGAGATTCGAGAGA 224
    |||||
Db 110 ProGluGluGlyGluValSerAspSerAspSerAsnSerSerSerSerSerSerSer 129
    :|||
QY 225 ---GCAACCGGACATATACAGCTTTTGCAGCAGCCTATTTCAGAGCCTCTA---- 276
    |||||
Db 130 AspSerSerSerGluAspGluGluPheHisAsp-GlyTyrglyGluAspLysMetGlyAs 149
    :|||
QY 277 -----GAGAAAGAGAGAGAAACCAACTT 299
    |||||
Db 149 pGluGluAspArgAlaArgLysGluGluMetThrGluLysGluArgGluGluGluLeuPhe 169
    :|||
QY 300 TGATCCAGCAGATGGGGAGTAAGTAGAGACCGCTTCTATACAATCATGCAATTCGA 359
    |||||
Db 169 eAsnArgileGluLysArgGluValLeuLysArgPheGluLysLysLysLysLysLeuLys 189
    :|||
QY 360 GGAGCAAGAACCCACCTGAGAAAGATCCCTAAACAAGAGAGTCTCAGATATCTGGGAA 419
    |||||
Db 189 sThrAlaLysLysLysGluLysLysGluLysLysGluGluGluGluGluGluGlu 209
    :|||
QY 420 GGAGGAGAGACATCAGTCACCATCTTAGACTCTTCGAGGAGAGATAAGGAA----- 471
    |||||
Db 209 sLysLysLeuThrGlnIleGluGluSerGlnValThrSerHisAsnLysGluArgArgSe 229
    :|||
QY 472 ----AAAGAAGAGTTGCTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 527
    |||||
Db 229 rlysgAspGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 249
    :|||
```

```
QY 528 AGAGGAG-----GCAAAGAAA-----ATGAAACAAA 554
    |||||
Db 249 gGluysArgLysAsnArgThrAlaGluLeuLeuAlaLysLysGlnProLeuLysThrSe 269
    :|||
QY 555 TAGTCTTCAAAATGAGGAAAGAGGAAACAAAGTGGAGC----- 595
    :|||
Db 269 rGluValTySerAspGluGluGluGluGlu-AspAspLysSerSerGluLysSerA 289
    :|||
QY 596 ----ACTGGTTTACCTCCAGGAAACATGAAATAATCCAAA----- 634
    |||||
Db 289 spArgSerSerArgThrSerSerSerAspGluGluGluGluGluGluGluGluProL 309
    :|||
QY 635 -----TCCATCAACCTTCTTATTAATATGTCATTTCTTCTCGAGGAGGAAAGA 680
    |||||
Db 309 ysSerGlnProValSerLeuProGluGluLeuAsnArgValArgLeuSerArgHisLysL 329
    :|||
QY 681 TTTGATGTTGGAATAACATTCCTTACTGTTGTGGAATACTGTCTCATGAGCATTTGTTTA 740
    :|||
Db 329 eu-----GluArgTrpCysHisMetProPhePheA 339
    :|||
QY 741 ATRAGCATACCATGAAACATGCCACTTGAAGATTCTCTGAGATCATGAGTTGTTTAC 800
    |||||
Db 339 laLys---ThrValThrGlyCysPheValArgIleGlyIleLysAsn-----H 354
    :|||
QY 801 ACTTGTCTCAAGCCTATCTATAGAGACCTTTGGATTAGAAATTATAGAACTAAAGTATCT 860
    |||||
Db 354 isAsnSer-LysProValTyArg-----ValAla 363
    :|||
QY 861 GAGATTACAGAGATCTCAGAGGTTATGTTCTTAACATATATATCAATG 908
    |||||
Db 364 GluileThrGlyValValGluThr-----AlaLysValTyGlnLeu 377
    :|||

RESULT 30
US-09-934-455-450
; Sequence 450, Application US/09934455
; Publication No. US20030121070A1
; GENERAL INFORMATION:
; APPLICANT: Creelman, Robert
; APPLICANT: Dubell, Arnold
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddie, James
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Reuber, Lynne
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Pineda, Omaiza
; TITLE OF INVENTION: Genes for Modifying Plant Traits IV
; FILE REFERENCE: MBI-0025
; CURRENT APPLICATION NUMBER: US/09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227439
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: MBI-0022
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: MBI-0023
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 516
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 450
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-934-455-450

Alignment Scores:
Pred. No.: 1.01 Length: 775
Score: 99.00 Matches: 45
Percent Similarity: 42.93% Conservative: 34
Best Local Similarity: 24.46% Mismatches: 54
Query Match: 6.03% Indels: 51
```

```
DB: 10 Gaps: 8
AF334735 (1-954) x US-09-934-455-450 (1-775)
QY 76 GGAACCGCGGCACACAGCTCGAGAGAAAGAGGTTCCATAGGCAGTCTTACCAAGAG 135
Db 153 GlySerGlyArgAspThrAlaGlnGluGluGluPro----- 165
QY 136 ATGTGATTCATTCCTCCACACCCACTACCAAGATTTCCCAAGGATTTGGG----- 186
Db 166 -----ArgThrProGluAsnValGlyLysSerAsn 175
QY 187 -----AATCTCTTGAAGGCTGACACCGGAGATTCTGAGAGCAACCGGACATATA 240
Db 176 GlyArgLysArgLeuGluLysThrThrProGluIleValSerAlaSerProAlaAsn--- 194
QY 241 CCAGCTTTTGCAGCAGCCTATTGTGAGACCTTCTAGAGAAAGAGAGAAACCACTTT 300
Db 195 ---SerMetAlaTrpAspTyrPhe-----PheMetValGluAsnMetProGlyProAsnLeu 212
QY 301 GATCCAGCAGAAATGGGGAGTAGGTAGAGACCGCTTC-----TATAACAAT 348
Db 213 AspAspArgGluValArgAsnGlyTyrGluAsnGlnSerSerHisPheGlnPheAsnGlu 232
QY 349 CATGCATTCGAGAGCAAGAACACCTGAGAAAGTGATCTTAACAAAGAGAGTCTCAG 408
Db 233 GluAspAspGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 252
QY 409 ATATCTGGGAG-----GAGGAGAGACATCAGTCACCATCTTAGACTCT 453
Db 253 ---SerGlyLysValValGluGluMetGluProLysThrProGluLysValGluGluGlu 271
QY 454 TCTGAGGAAGATAAGAAAAAGAGAGTTGCTGCTGCTCAAAATCCAAAGCTGCTCCGG 513
Db 272 GluGluGluAspGluGluAspGlu----- 280
QY 514 GGACACATAGCCAGAGAGAGGAGGAAAGAAATGAAATAGTCTTCAAAATAGGAA 573
Db 281 -----GluGluGluGluGluGluGluGluGluValValGluValGluValGluVal 296
QY 574 AAGAGGAAA 583
Db 296 slyLysLys 599

RESULT 31
US-10-437-963-183719
; Sequence 183719, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 183719
; LENGTH: 1537
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_80783C.1.pap
US-10-437-963-183719
Alignment Scores: 1.21 Length: 1537
Pred. No.: 1537
```

```
Score: 99.00 Matches: 45
Percent Similarity: 39.89% Conservative: 26
Best Local Similarity: 25.28% Mismatches: 66
Query Match: 6.03% Indels: 41
DB: 16 Gaps: 7
AF334735 (1-954) x US-10-437-963-183719 (1-1537)
QY 60 AACTAGAAAAACAACCGGACCGGCGGCACAGCTCGGAGAGAAAGAGGTTCCATAGGC 119
Db 146 AsnHisGluGluAsnGlyGluLysHisGluGluAsnGlyGluLysGlnAsnGluAsnGly 165
QY 120 AGTCTTACCAGAAGATGTGATTCATTTCTTCCAAACCCACTCCGCAATTTCCCAAGG 179
Db 166 GluLysCysGluGluAsnGlyGlyLysProAlaGln-----IleSerAspAspGlnAla 183
QY 180 ATTTGGGAATCTTCTTGAAGGCTGACACCGGAGATTCTGAGAGACCAACCGGACATAT 239
Db 184 ValAsnGlyCysAsnGlySerAlaAspLysLysAspLysGluGluGluThr----- 200
QY 240 ACCAGCTTTTGCAGCAGCCTATTGTGAGAGCCTTCTAGAGAAAGAGAGAAACCAACTT 299
Db 201 -----GluLysAspAsnGluLys-GluLys----- 208
QY 300 TGATCCAGCAGAAATCGGGAGTAAAGTAGAACCGCTTCTATAACAATCATGCAATCGA 359
Db 209 -----GluAspLys-----GluGluGluThrG 216
QY 360 GGAGCAAGAACCCACCTGAGAAAGTGATCTTAACAAGAGAGTCTCAGATATCTGGGAA 419
Db 216 uLysAspAsnGluLysGluLysGluAsp---LysGluGluGluThrLysLysAspAsnG 235
QY 420 GGAGGAAGAGACATCAGTCACCATCTTAGACTCTTCTGAGGAAGATAAGAAAAAGAA-- 477
Db 235 uLysGluLysGluGlnLeuMetGlyThrAspGluLysGluLysGluLysGluLysGlu 255
QY 478 -GAGTTTGTGTGTCAAAATCCAAAGCTGCTTCCGGGGACACATAGCCAGAGAGAGGC 536
Db 255 pGluAsnGluGluGluLysLeuGlu-----GluGluG 266
QY 537 AAGAAATGAACAACAATAGTCTTCAAAATGAGGAGAAAGAGGAAAGAAACAAG 588
Db 266 uLysLysAspLysGluGluLysLeuGluLysGluLysGluLysGluLysGluGlu 283

RESULT 32
US-10-087-192-1197
; Sequence 1197, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1197
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-087-192-1197
Alignment Scores: 0.961 Length: 426
Pred. No.: 98.50 Matches: 41
Percent Similarity: 41.26% Conservative: 18
Best Local Similarity: 28.67% Mismatches: 53
```

```

Query Match: 6.00% Indels: 31
DB: 13 Gaps: 7
AF334735 (1-954) x US-10-087-192-1197 (1-426)

Qy 40 AGAAAAACGGTTACCCGCACTAGAAAAACCAACCGACCGCGGCACCGACTCGGAG 99
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4 ArgSerLeuGlySerArgAlaalaargAspArgThrAlaAspArgGly----- 19

Qy 100 AGAAAGGAGTTCCATAGSCA---GTTCTTACCAAGAAGATGTCGATTCCATTCTCCAAC 156
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 20 ArgGlyArgArgProGlyAlaSerAlaCysThrGlyArgMetSerIle----- 35

Qy 157 ACCCACTACCGAATTCCACAAGATTTGGGAATCTTCTTGAAGGGCTGACACCGAGATT 216
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36 -----GluIleProAlaGlyLeuThrGluLeuLeuGlnGlyPheThrValGluVal 52

Qy 217 CTGAGAGACACCGGACAAATATACCACTTTTGCAGCAGCCTATTTTGAGAGCCTTCTA 276
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 53 LeuArgHisGlnProAlaAspLeuLeuGluPheAlaLeuGlnHisPheThrArgLeuGln 72

Qy 277 GAGAAAGAGAGAAAACC-----AACTTGATCCAGCAGAATGGGGAGT 321
   : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 73 GlnGluAsnGluArgLysGlyAlaalaArgPheGlyHisGluGlyArgThrTrpGly--- 91

Qy 322 AAGGTAGAACCGCTTCTATAACAATCATGCAATTCGAGGAGCAAGAACCCATCGAGAA 381
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 92 -----AspAlaGlyAlaalaalaGlyGlyThrProSerLys 104

Qy 382 -----AGTGATCTTAAACAAGAGAGTCTCAGATATCTGGGAAGGAGAA 426
   : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 105 GlyValAsnPheAlaGluGluProMetArgSerAspSerGlu---AsnGlyGluGluGlu 123

Qy 427 GAGACATCA 435
   ||| ||| : : :
Db 124 GluAlaAla 126

RESULT 33
US-10-424-599-166145
; Sequence 166145, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 166145
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(273)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MBT3847_121044C.1.pep
US-10-424-599-166145

Alignment Scores:
Pred. No.: 0.955 Length: 273
Score: 98.00 Matches: 43
Percent Similarity: 38.15% Conservative: 23
Best Local Similarity: 24.86% Mismatches: 39
Query Match: 5.97% Indels: 68
DB: 15 Gaps: 10
AF334735 (1-954) x US-10-424-599-166145 (1-273)

```


Db 721 ysGluGluLysProLysAspValProGlu---LysLysLysAlaGluSerProValLysG 740
:::|::|::|
QY 344 ACAATCATGCTTCAG-----GAGCAAG 367
|::|
Db 740 LuGluAlaValAlaGluValThrIleThrLysSerValLysValHisLeuGluLysG 760
|::|
QY 368 ACCACCTGAGAAAGTATCTTAACAAGAGAGTCTCAG-----A 409
|::|
Db 760 LuThrLysGluGluGlyLysProLeuGlnGlnGluLysGluLysAlaGlyGlyG 780
|::|
QY 410 TATCTGGGAGGAGGAGAGACATCAGTCACCATCTTAGACTCTTCTCAGGAGAGATAAGG 469
|::|
Db 780 LuGlyGlySerGluGluGluGlySer-----AspLysGlyAlaLysGlySerA 796
|::|
QY 470 AAAAGAGAGAGTCTGCTGCT-----GTCAAAATCCAAG 502
|::|
Db 796 rgLysGluAspIleAlaValaenGlyGluValGluGlyLysGluValGluGlnGluT 816
|::|
QY 503 CTGCTCTCCGGGACACATAGCCAGAGGAGGCAAGAAATGAAACAAATAGTCTTC 562
|::|
Db 816 hrLysGluLysGlySer---GlyArgGluGluGluLysGlyValThrAenGlyLeuA 835
|::|
QY 563 AAAATCAGGAAAAAGAGGAAAAACAAGTCAAGG 594
|::|
Db 835 spLeuSerProAlaAspGluLysLysGlyGly 845
|::|

RESULT 37

US-10-424-599-174175
; Sequence 174175, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 174175
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_12829C.1.psp
US-10-424-599-174175

Alignment Scores:
Pred. No.: 0.981 Length: 197
Score: 97.50 Matches: 45
Percent Similarity: 38.89% Conservative: 32
Best Local Similarity: 22.73% Mismatches: 68
Query Match: 5.94% Indels: 54
DB: 15 Gaps: 7

AF334735 (1-954) x US-10-424-599-174175 (1-197)

QY 32 TTTTAAAGAAAAA-----ACGTTTACCAGCAACTAGAAAAACAACCG 76
|::|
Db 37 PheThrLysGluGlnLeuArgValGlnValThrSerThrProValLeuArgIleAsnGly 56
|::|
QY 77 GAACCGGGCCACAGCTCGGAGAGAAAGAGGTTCCATAGCAGTTC---TTACCAAGA 133
|::|
Db 57 GluArgGlnIleValGluAsnLysArgArgPheSerArgGluPheSerIleProPro 76
|::|
QY 134 AGATGT-----CGATTCCATTCTCCAAACACCCACTACCAATTC 172
|::|
Db 77 TyrCysAspThrAsnAspValSerAlaLysPheGluGlyGlyValLeuSerIleLysPhe 96
|::|
QY 173 CACAAGGATTTGGGAATCTTCTTGAAGGGCTGACACGGGAGATTCTCGAGAGAGCAACGG 232
|::|

Db 97 ProLys----- 98
|::|
QY 233 ACAATATACAGCTTTTGCAGCAGCCTATTTTTCAGAGCCTTCTAGAGAAAAGAGAGAAA 292
|::|
Db 99 -----LeuIleThrProAlaAlaArgSerGlnProGlnProGlnGluAla 113
|::|
QY 293 CCAACTTGTATCCAGCAGATGG---GGGAGTAAGGTAGAGACCGCTTCTATACCAAT 348
|::|
Db 114 ProThrMetProGlnLysGluLysGluProSerGlnGlnAspGlnValHisLysGln 133
|::|
QY 349 CATGCAATTCGAGGAGCAAGAACCACTGAGAAAAAGTGATCCTTAAACAAGAA---GAGTCT 405
|::|
Db 134 GluSerLeuGlnLysGluLysGluProIleThrSerAspGluLysGluLysGluAsnLysThr 153
|::|
QY 406 CAGATATCTCGGAGGAGGAGAGACATCAGTCACCATCTTAGACTCTTCTTGAGAGAAAT 465
|::|
Db 154 GluGluSerSerProGlnLysGluAsnGluProIleSerAspAspLysGluLysAsp 173
|::|
QY 466 AAGNAAAAGAGAGGTGCTGCTGCTCAAAATCCAAAGCTGCTTCCGGGGACACATAGCC 525
|::|
Db 174 AsnLysThrGluGluValValAla----- 181
|::|
QY 526 AGAGAGGAGGCAAAATAATGAAACAATAATAGTCTTCAAAATGAGGAAAAAGAG 579
|::|
Db 182 -----GluLysLysValArgThrAsnGlyValProGluThrAlaLysGlu 196
|::|

RESULT 38

US-10-437-963-202437
; Sequence 202437, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 202437
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_97717C.1.psp
US-10-437-963-202437

Alignment Scores:
Pred. No.: 1.31 Length: 574
Score: 97.50 Matches: 26
Percent Similarity: 57.53% Conservative: 16
Best Local Similarity: 35.62% Mismatches: 28
Query Match: 5.94% Indels: 3
DB: 16 Gaps: 2

AF334735 (1-954) x US-10-437-963-202437 (1-574)

QY 340 TATAAATCATCATGCTTCGAGAGCAAGAACCCACTGAGAAAAGTATCTTAAACAAGAA 399
|::|
Db 62 HisAsnAsnGluAlaValGlnGluValGlyArgGlyGluAsnSerSerLeuGlnGlyGlu 81
|::|
QY 400 -----GAGTCTCAGATATCTGGGAGGAGAGAGACATCAGTCACCATCTTAGACTCT 453
|::|
Db 82 ValValValArgAspValSerGlnAspLeuGluLysGlnAsnThrValValSerAspAla 101
|::|
QY 454 TCTGAGGAA---GATAAGGAAAAAGAGAGGTTGCTGTCTGTCACAAATCCAAAGCTGCCTTC 510
|::|

Db 102 SerAspProGluArgLeuArgGluGluGlnAlaAlaValysAlaGlnAlaAlaPhe 121
QY 511 CGGGGACACATACCGACAGAGAGGAGGCAAAAGAAATGAAA 549
Db 122 ArgGlyTyrluAlaAlaArgAlaPheArgAlaLeuLys 134

RESULT 39

US-10-437-963-158544
; Sequence 158544, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 158544
; LENGTH: 620
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_58008C.1.pap
US-10-437-963-158544

Alignment Scores:
Pred. No.: 1.34 Length: 620
Score: 97.50 Matches: 51
Percent Similarity: 40.85% Conservative: 36
Best Local Similarity: 23.94% Mismatches: 74
Query Match: 5.94% Indels: 52
DB: 16 Gaps: 11

AF334735 (1-954) x US-10-437-963-158544 (1-620)

QY 75 CGGAACCGGGCGCAC-----CAGCTCGGAGAGAAAGGAGGTTCCATAGGCAGT 122
Db 93 ArgAsnArgArgAlaTyrluAspArgLeuValArgGluGlyGluTyrluPheSerGlu 112
QY 123 TCTTACCAAGAGATGTCGATTCATTCCTCCACACCCACTACCGAATTCACAGGATT 182
Db 113 GluAlaMetArgGluArgGluProTyrluLeuHisGluTyrluGly-----ArgPhe 130
QY 183 TGGGAATCTTGTGAAGGCTGACGCGAGATTCGAGAGACCAACCGACATATA-- 240
Db 131 GlnAspProthrglyArgAlaMetAlaArgProGlyGluArgTrpSer-GluThrLeuMe 150
QY 241 ----CCAGCTTTTTCGACGACCTATTTCGAGACCTTCCTAGAGACCTTCAGAGAAAGAGAGAAACCAA 296
Db 150 tArgArgAlaGluGluAlaValIleValGluLysIleArgGlyGluGlnIleArgGlu 170
QY 297 CTTTGTATCCACAGAAATGG---GGGAGTAGGTAGAGACCGCTTCTATAACAAATCATGC 353
Db 170 yValAspProSerGluTrpValGlyGlyGlyAlaGluGlu-----Al 184
QY 354 ATTCTGAGGACAGAACCCCTGAGAAAGTGTATCTTAACAGAGAGAGTCTCAGATATC 413
Db 184 aMetGluGluGlnGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 204
QY 414 TGGGAGGAGGAGAG----- 429
Db 204 uSerGluGluGluGlnGlnGlyMetAspMetGluGluLysGlyArgGlyAlaAspAsnSerSe 224
QY 430 -----ACATGAGTCACC---ATCTTAGACTCT----- 453

Db 224 rGluAsnProSerValThrGluValValAspSerAsnLysThrAlaGlyAlaSerLysGln 244
QY 454 -----TCTGAGGAAGATAGGAAAAAGAA---CAGGTTGCTGCTGTCAAAATCCAAAGC 503
Db 244 nThrLeuSerAlaGluAspMetGluAspGlnLeuGluGlnPheThrSerLeuMetGlnGln 264
QY 504 TGCCTTCGGGACACATAGCCAGAGAGAGGCAAGAAAGAAATGAAACAAATAGTCTTCA 563
Db 264 nLysPhe-----LeuSerGlyGluAspSerGluHisMetAspTyrSerArgIleAs 281
QY 564 AAATCAGGAAAGAGAGAAACCAAGTGAGGACACTGG 600
Db 281 pAsnAspGluMetLeuAspAsp-----HisTrp 290

RESULT 40

US-10-092-900A-20
; Sequence 20, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Liu, Yi
; APPLICANT: Anderson, David W.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Catterton, Elina
; APPLICANT: Leite, Mario W.
; APPLICANT: Zhong, Hainong
; APPLICANT: Alsobrook, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-290C
; CURRENT APPLICATION NUMBER: US/10/092,900A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: USSN 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/283,675
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: USSN 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: USSN 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/274,191
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: USSN 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: USSN 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: USSN 60/294,899
; PRIOR FILING DATE: 2001-05-31

```
; PRIOR APPLICATION NUMBER: USSN 60/287,424
; PRIOR FILING DATE: 2001-04-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 768
; SEQ ID NO 20
; LENGTH: 749
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-900A-20

Alignment Scores:
Pred. No.: 1.41 Length: 749
Score: 97.50 Matches: 50
Percent Similarity: 38.58% Conservative: 48
Best Local Similarity: 19.69% Mismatches: 63
Query Match: 5.94% Indels: 93
DB: 15 Gaps: 11

AF334735 (1-954) x US-10-092-900A-20 (1-749)
QY 70 ACAACCGGAACCGCGGCACC-----AGCTCGGAG 99
Db SerThrGlySerGlyAsnThrGluHisSerCysSerSerGlnLysGlnLeuSerIleGln 203
QY 100 AGAAGAGGAGTCCATAGGCACTTTACCAAGAGAGATGTCGATTCCTCCACACC 159
Db HisArgGlnThrGlnSerAspLeuThrIleGluLysIleSerAlaLeuGluAsnSerLys 223
QY 160 CACTACCGAATTCACAGAGATTGGGAATCTTCTGAAGGCGTCACACGCGAGATT--- 216
Db 224 AsnSerAspLeuGluLysLysGluGlyArg---IleAspAspLeuArgAlaIleCys 242
QY 217 ---CTGAGAGACAAACCGGACCAATATACAGCTTTTGGCAGCAGCCTATTTGAGAGC--- 270
Db 243 AspLeuArgArgGlnIleAepGluGlnGlnLysMetLeuGluLysTyrlLysGluArgLeu 262
QY 271 -----CTTCTAGAGAAA----- 282
Db 263 AsnArgCysValThrMetSerLysLysLeuLeuIleGluLysSerLysGlnGluLysMet 282
QY 283 -----AGAGAGAAAACC-----AACTTTGATCCAGCAAGAA 312
Db 283 AlaCysArgAspLysSerMetGlnAspArgLeuArgLeuGlyHisPheThrThrSerAsp 302
QY 313 TGGGGAGTAAGGTAGAGACCGCTTCTATAACAATCATGCAATTCGAGAG----- 363
Db 303 HisGlyAlaLysPheThrGluGlnThrAspGlyTyrlAlaPheGlnAsnLeuIleLys 322
QY 363 ----- 363
Db 323 GlnGlnGluArgIleAsnSerGlnArgGluGluIleGluArgGlnArgLysMetLeuAla 342
QY 364 -----CAAGAACCCACTCTGAGAAAAGTGCCTTAACAAGAA 399
Db 343 LysArgLysProProAlaMetGlyGlnAlaProProAlaThrAsnGluGlnLysGlnTrp 362
QY 400 GAGTCTCAGATATCTGGGAGGAGGAGAGACATCAGTCACCATCTTAGACTCTTCTGAG 459
Db 363 LysSerLysThrAsnGlyAlaGluAsnGluThr---LeuThrLeu-----Lys 377
QY 460 GAAGATAAGGAAAGAGAGAGGTTGCTGCTGTCAAAATCCAGTGCCTTCGGGGAGAC 519
Db 378 GluTyrlHisGluGlnGluGluIlePheLysLeuArgLeu-----GlyHis 392
QY 520 ATAGCCAGAGAGGAGCA-----AAGAAA 543
Db 393 LeuLysLysGluGluAlaGluIleGlnAlaGluLeuGluArgLeuGluArgValArgLys 412
QY 544 ATGAAAACAAATAGTCTTCAAAATGAGGAAAAGAGAAAC 585
Db 413 LeuHisIleArgGluValLysArgIleHisAsnGluAspAsn 426
```

RESULT 41

```
US-10-369-493-6247
; Sequence 6247, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 6247
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-6247

Alignment Scores:
Pred. No.: 1.41 Length: 484
Score: 97.00 Matches: 55
Percent Similarity: 41.36% Conservative: 24
Best Local Similarity: 28.80% Mismatches: 66
Query Match: 5.77% Indels: 46
DB: 15 Gaps: 8

AF334735 (1-954) x US-10-369-493-6247 (1-484)
QY 803 AGTGTAACAACATCATGATCTCAGAGAAATCTTCAAGTGCATGTTTCAATGTGATGCT 744
Db 275 SerGluAsnArgAlaThrIleGluGlnSerSer-AlaCys---AsnArgLeuLe 293
QY 743 TATTAACAATGTCATGACAGATTTTCACACAGTAACCAAGTGTATTTCACACATC 684
Db 293 uLeuGlySerIleLeuLeuAsn-----Phe-ArgHisA 304
QY 683 AAATCTTCTCTCCCTCAGGAAGAAATGACATTAATAAGAAGTTGATGGATTGATTATT 624
Db 304 snLeuLeuLeuTrpLeuArgLeuLeuLeuLeuLeuLeuSerPheLeuSerLeuPheLeuP 324
QY 623 TTTCATGTTTCTCGAGGTAAACACAGTGCTCTCACTTGTTTCTCTCTTTTCTCTCATTT 564
Db 324 he--LeuLeuLeuGlyLeuLeuLeuSerSerLeuValLeuLeuSerSerValIleL 343
QY 563 TGAAGACTATTGTTGTTTTCATTTCTTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 504
Db 343 euLeuLeuPheIleIleSerThrIleLeuLeuLeuLeu-----Phe-ArgHisA 355
QY 503 GCTTGGATTTTGACAGCAGCAACCTCTTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCT 444
Db 356 -----PhePheLeuPheLeuPheLeu----- 363
QY 443 ATGCTGACTGATGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 384
Db 364 -----LeuPheLeuLeuLeuPheLeu-----LeuLeuLeuLeuAsnIleP 378
QY 383 CTTTCTCAGTGGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 336
Db 378 heLeuLeuArgGlyPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 398
QY 335 CGGTCTTCTACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 288
Db 398 euLeuPhePheLeuAspTyrlArgLeuCysTrpMetHisGlnSerProArgLysArgLysL 418
QY 287 TCTCTTTTCTCTAGAGGCTCTCAAAATA 259
Db 418 euLysAlaLeuLysCysLeuLysVal 427
```


185	GlubArgtysAsnValLysGlyIleAArgAspAspIleGluGluGluAspAspGlnGluAla	204
Db		
508	TTCCGGGGACACATAGCCAGAGAGGAGGCAAGAAAATGAAAACAAATAGTCTTCAAAAT	567
Qy		
205	TyrPheArgTyrMetAlaGluAsn-----ProThrAlaGlyValValGln	219
Db		
568	GAGGAAAAAGAGGAAAAAC	585
Qy		
220	GluGluGluGluAspAsn	225
Db		

RESULT 45
US-10-017-161-2198
; Sequence 2198, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1

```

? LENGTH: 430
? TYPE: PRNT
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: MOD_RES
? LOCATION: (151)
? OTHER INFORMATION: Variable amino acid
? FEATURE:
? NAME/KEY: MOD_RES
? LOCATION: (354)
? OTHER INFORMATION: Variable amino acid
? FEATURE:
? NAME/KEY: MOD_RES
? LOCATION: (385)..(418)
? OTHER INFORMATION: Variable amino acid
? US-10-017-161-2198

```

```

; OTHER INFORMATION: Variable amino acid
; FEATURE:
; CDS INFORMATION: Validated amino acid
; NAME/KEY: MOD RES
; LOCATION: (385)..(418)
; OTHER INFORMATION: Variable amino acid
US-10-017-161-2198

Alignment Scores:
Pred. No.: 1.54 Length: 438
Score: 96.50 Matches: 58
Percent Similarity: 37.12% Conservative: 27
Best Local Similarity: 25.33% Mismatches: 48
Query Match: 5.74% Indels: 96
DB: 14 Gaps: 12

AF334735 (1-954) x US-10-017-161-2198 (1-438)

Qy 600 CCAGTGCCCTCA-----CTTGTTTTCTCTTTTTCTCAATTGGAAGACTATTGTTTT 547
Db 200 ProLeuSerSerLeuLeuLeuPheLeuLeuPhePheLeuLeuHisSerLeuPhePhe 219
Qy 546 CATTTT-----CTTTGCCCTCTCTCTGGCTATGTGTCCCG 511
Db 220 SerPhePheAanAanHisSerProSerLeuProLeuProProPheSerPro 239
Qy 510 GAAGGCAGCTGGATTGTGACAGCAGCAACTCTCTCTTTTC----- 469
Db 240 SerPhePheLeuPhePhePhePheSerPhePhePhePhePhePhePheHis 259
Qy 468 -----CTTATCTCTCCTCAGAAGAGCTCTAAGATGGTGACTGATGCTCTTC----- 424
Db 260 PheSerProLeuLeuPheLe-----PheLeuPheSerPhe 271
Qy 423 -----CTCCCTCCAGATATCTGAGACTCTTCTTGTTTAGATACACTTTTCTCAGGTGG 370
Db 423 -----CTCCCTCCAGATATCTGAGACTCTTCTTGTTTAGATACACTTTTCTCAGGTGG 370

```



```

Best Local Similarity: 24.07%      Mismatches: 59
Query Match: 5.85%                Indels: 0
DB: 13                            Gaps: 0

AF334735 (1-954) x US-10-101-487-81 (1-191)

Qy 277 GAGAAAGAGAGAAAACCACTTTGATCCAGCAGAAATGGGGAGTAAGTAGAAGACCOC 336
    |||::|||:::   ::|||::|||:::   ::|||:::
Db 77 GluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGlu 96

Qy 337 TTTATAACAATCATGCATTTCGAGGACGAAGAACCACCTGAGAAAAGTGATCCTAAACA 396
    |||::|||:::   |||::|||:::   |||:::   ::|||:::
Db 97 GluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGlu 116

Qy 397 GAAGAGTCTCAGATATCTGGGAAGGAGGAGACATCAGTCACCATCTTAGACTCTTCT 456
    |||::|||:::   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db 117 GluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGlu 136

Qy 457 GAGGAAGATAAGGAAAAAGAGAGGTGCTGCTGTCAAATGCCAAGTGCCTTCGGGGA 516
    |||::|||:::   |||::|||::|||::|||::|||::|||::|||::|||::|||::
Db 137 GluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGlu 156

Qy 517 CACATGCCAGAGGAGGCAAGAAATGAACAATAGTCTTCAAATGAGGAAAAA 576
    |||::|||:::   |||::|||::|||::|||::|||::|||::|||::|||::|||::
Db 157 GluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGlu 176

Qy 577 GAGGAAACAAGTCGAGGACACTGG 600
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db 177 GluGlUGluCysThrHisTrp 184

RESULT 49
US-10-101-487-42
; Sequence 42, Application US/10101487
; Publication No. US20020169125A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, DAVID W.
; APPLICANT: BERGMAN, PHILIP A.
; APPLICANT: LOFQUIST, ALAN
; APPLICANT: PIETZ, GREGORY E.
; APPLICANT: TOMPKINS, CHRISTOPHER K.
; APPLICANT: WAGONER JR., DAVID W.
; TITLE OF INVENTION: RECOMBINANT PRODUCTION OF POLYANIONIC POLYMERS AND USES
; FILE REFERENCE: 077319/0329
; CURRENT APPLICATION NUMBER: US/101,487
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/277,705
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic fusion
US-10-101-487-42

Alignment Scores:
Pred. No.: 1.39                      Length: 198
Score: 96.00                        Matches: 26
Percent Similarity: 45.37%          Conservative: 23
Best Local Similarity: 24.07%       Mismatches: 59
Query Match: 5.85%                  Indels: 0
DB: 13                              Gaps: 0

AF334735 (1-954) x US-10-101-487-42 (1-198)

Qy 277 GAGAAAGAGAGAAAACCACTTTGATCCAGCAGAAATGGGGAGTAAGTAGAAGACCOC 336
    |||::|||:::   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db 84 GluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGluGlUGlu 103

```

QY 337 TTCTATAACATCATCGATTTCAGGAGCAAGAACACCTGAGAAAAGTGTATCTTAACAA 396
Db 104 GluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 123
QY 397 GAAGAGTCTCAGATATCTGGGAGGAGGAGAGACATCAGTCACCATCTTAGACTCTTCT 456
Db 124 GluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 143
QY 457 GAGGAAGATAGGAAAAAGAGGTTGCTGTGTCAAAATCCAAAGTCCTTCGCGGGA 516
Db 144 GluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 163
QY 517 CACATAGCCAGAGAGAGCAAGAAAATGAAACAAATAGTCTTCAAAATGAGGAAAA 576
Db 164 GluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 183
QY 577 GAGGAAAAACAGTGAGGACACTGG 600
Db 184 GluGluGluCysThrThrHisTrp 191

RESULT 50

US-10-273-334-34
; Sequence 34, Application US/10273334
; Publication No. US20030129631A1
; GENERAL INFORMATION:
; APPLICANT: Pasternack, Gary R.
; APPLICANT: Kocheavar, Gerald J.
; APPLICANT: Brody, Jonathan R.
; APPLICANT: Kodkol, Shrihari S.
; TITLE OF INVENTION: GENE FAMILY WITH TRANSFORMATION MODULATING ACTIVITY
; FILE REFERENCE: 031787.0076
; CURRENT APPLICATION NUMBER: US/10/273,334
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: US/09/591,500
; PRIOR FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: PCT/US98/26433
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 60/069,677
; PRIOR FILING DATE: 1997-12-11
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-273-334-34

Alignment Scores:
Pred. No.: 1.48 Length: 249
Score: 96.00 Matches: 35
Percent Similarity: 42.42% Conservative: 21
Best Local Similarity: 26.52% Mismatches: 49
Query Match: 5.85% Indels: 27
DB: 14 Gaps: 4

AF334735 (1-954) x US-10-273-334-34 (1-249)

QY 193 CTTGAAGGCTGACACGCGAGATTCTGAGAGAGCAACCGACAATATACAGCTTTTGCA 252
Db 145 LeuAspGlyCysAspArgAsp---LysGluAlaProAsp-----SerAspAla 160
QY 253 GCAGCCTATTTTGAAGCCTTCTAGAGAAAAGAGAGAAACCAACTTTTGATCCAGAGAA 312
Db 161 GluGlyTyrrValGluGlyLeuAspAspGluGluGluAspGluAsp----- 175
QY 313 TGGGGGAGTAGAGACCGCTTCTATACATCATGCAATTCGAGGAGCAAGAACCA 372
Db 176 -----GluGluGlyTyrrAspGluAspAlaGlnValValAspGluGlu 190
QY 373 CCTGAGAAAAGTGATCCTAAACAAGAGAGTCTCAGATATCTGGGAGGAGGAGAGACA 432
Db 191 AspGluAspGluGluGluGlyGluGluGluAspValserGlyGluGluGluGluAsp 210

QY 433 TCAGTCACCATCTTAGACTCTTCTGAGGAGAGATAGGAAAAAGAGAGGTTGCTGCTGTC 492
Db 211 GluGluGlyTyrrAsnAspGlyGluValAspAspGluGluAspGluGlu----- 226
QY 493 AAAATCCAGCTGCCTTCGCGGAGCACATAGCCAGAGAGGAGGCAAGAAAATGAAAAACA 552
Db 227 -----GluLeuGlyGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 238
QY 553 AATAGTCTTCAAAATGAGGAAAAAGAGGAAAAACA 586
Db 239 GluAsnLeu-LysMetArgGluLysMetMetThr 249

Search completed: February 15, 2005, 10:28:54
Job time : 174 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 15, 2005, 09:52:34 ; Search time 32.5 Seconds
(without alignments)
5648.664 Million cell updates/sec

Title: AF334735
Perfect score: 1641
Sequence: 1 TCGCCCTTCTCGCCGCG.....AAAAAAAAAAAAAAAAAAAA 954

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 56832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q/cn2.1/USPTO.spool_p/AF334735/runat_15022005.094943.10099/app_query.fasta.1.1095
-DB=PIR79 -OFMT=fasta -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=AF334735 @CGN 1.1 63 @runat_15022005.094943.10099 -NPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	778	47.4	151	2	I38243
2	569	34.7	146	2	I46506
3	558	34.0	149	2	S49527
4	134	8.2	189	2	T29159
5	118.5	7.2	404	1	OKHU2R
6	115.5	7.0	400	1	OKG02R
7	115	7.0	532	2	T06029
8	113.5	6.9	443	2	T48593
9	110	6.7	103	2	T45769
10	108	6.6	403	1	A43435
11	108	6.6	1132	2	S37932
12	107.5	6.6	805	2	E70474
13	106	6.5	1178	2	S78475
14	105.5	6.4	1351	2	C71607

15	104	6.3	304	2	S44897
16	104	6.3	699	2	T21379
17	103	6.3	238	2	I52638
18	102.5	6.2	204	2	S45705
19	102	6.2	213	2	J00075
20	101.5	6.2	1003	2	T24545
21	101	6.2	419	2	T18450
22	101	6.2	475	2	S52893
23	101	6.2	647	2	E64170
24	101	6.2	1658	2	S55101
25	100.5	6.1	491	2	C97267
26	100	6.1	205	2	C30491
27	100	6.1	762	2	G88436
28	100	6.1	783	2	F84514
29	100	6.1	791	2	T24435
30	99.5	6.1	414	2	A39205
31	99.5	6.1	689	2	T34025
32	99.5	6.1	858	2	S15762
33	99	6.0	189	2	T02570
34	99	6.0	241	2	S44893
35	99	6.0	644	2	S55395
36	99	6.0	1192	2	A71623
37	98.5	6.0	213	2	S45706
38	98.5	6.0	418	2	T15827
39	98.5	6.0	706	2	A45990
40	98	6.0	916	2	A27864
41	98	6.0	2453	2	S60254
42	97.5	5.9	1094	2	S49313
43	97	5.9	407	1	E0BEQ3
44	97	5.8	484	2	F88924
45	97	5.9	918	2	T38786

ALIGNMENTS

RESULT 1

I38243
zona pellucida binding protein Sp17 [similarity] - human
C:Species: Homo sapiens (man)
C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C:Accession: I38243; S52921
R:Richardson, R.T.; Yamashiki, N.; O'Rand, M.G.
Dev. Biol. 165, 688-701, 1994
A:Title: Sequence of a rabbit sperm zona pellucida binding protein and localization during fertilization
A:Reference number: I38243; MUID:95046885; PMID:7525387
A:Accession: I38243
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-151 <RES>
A:Cross-references: UNIPROT:Q15506; EMBL:Z48570; NID:G695580; PIDN:CAA88459.1; PID:G695580
R:Lea, I.A.; Richardson, R.T.; Widgren, E.E.; O'Rand, M.G.
submitted to the EMBL Data Library, March 1995
A:Description: Human Sp17: a sperm-zona binding protein.
A:Reference number: S52921
A:Accession: S52921
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-151 <LEA>
C:Superfamily: sperm surface protein Sp17

Alignment Scores:
Pred. No.: 4,568-64 Length: 151
Score: 778.00 Matches: 151
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 47.41% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x I38243 (1-151)

QY 136 ATGTGATTCATTTCTCCACACCCACTACCAAGATTCGAGATTCCTCTT 195
|||||
DB 1 MetSerIleProPheSerAsnThrHisTyArgGlieProGlnGlyPheGlyAsnLeuLeu 20

QY 196 GAAGGGCTGACCGGAGATTCTGAGAGCAACCGGCAATATATACCGCTTTTCAGCA 255
Db 21 GluGlyLeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAlaAla 40
QY 256 GCCTATTTTGTAGAGCCTTCTAGAGAAAAGAGAGAAAACCAACTTTGTATCCAGCAGAAATGG 315
Db 41 AlaTyrPheGluSerLeuLeuGluLysArgGluLysThrAsnPheAspProAlaGluTrp 60
QY 316 GGGAGTAAGGTAGAGACCGCTTCTATTAACAATCATGCAATTCGAGGAGCAAGAACCACT 375
Db 61 GlySerLysValGluAspArgPheTyrAsnAsnHisAlaPheGluGluGlnPro 80
QY 376 GAGAAAGTGTCTTAAACAGAGTCTCAGATATCTGGGAAAGGAGGAGAGACATCA 435
Db 81 GluLysSerAspProLysGlnGluSerGlnIleSerGlyLysGluGluGluThrSer 100
QY 436 GTCACCATCTTAGACTCTTCTGAGGAAGATAGGAAAAGAGAGCTTGTCTGTCAAA 495
Db 101 ValThrIleLeuAspSerSerGluGluAspLysGluLysGluLysValAlaAlaValLys 120
QY 496 ATCCAAGTCTGCTCCGGGACACATAGCCAGAGAGGCGCAAGAAAATGAAAACAAAT 555
Db 121 IleGlnAlaAlaPheArgGlyHisIleAlaArgGluGluAlaLysLysMetLysThrAsn 140
QY 556 AGTCTTCAAAATGAGGAAAAGAGGAAACAG 588
Db 141 SerLeuGlnAsnGluLysGluLysGluLys 151
RESULT 2
I46506
zona pellucida binding protein Sp17 [similarity] - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 14-Feb-1997 #sequence revision 14-Feb-1997 #text_change 09-Jul-2004
C:Accession: I46506; S58441; S31764
R:Richardson, R.T.; Yamasaki, N.; O'Rand, M.G.
Dev. Biol. 165, 688-701, 1994
A>Title: Sequence of a rabbit sperm zona pellucida binding protein and localization during spermatogenesis
A:Reference number: 138243; MUID:95046885; PMID:7525387
A:Accession: I46506
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-146 <RIC>
A:Cross-references: UNIPROT:P36425; EMBL:Z20655; NID:g479083; PIDN:CAA79674.1; PID:g479083
A:Accession: S58441
A>Status: preliminary
A:Molecule type: protein
A:Residues: 113-126 <RI3>
R:Richardson, R.T.; Yamasaki, N.; O'Rand, M.G.
submitted to the EMBL Data Library, January 1993
A:Description: Sequence of a mammalian sperm surface zona pellucida binding protein and localization during spermatogenesis
A:Reference number: S31764
A:Accession: S31764
A:Molecule type: DNA
A:Residues: 1-122, 'DT', <RI2>
A:Cross-references: EMBL:Z20655
C:Superfamily: sperm surface protein Sp17
Alignment Scores:
Pred. No.: 1,196-44 Length: 146
Score: 569.00 Matches: 109
Percent Similarity: 86.67% Conservative: 21
Best Local Similarity: 72.67% Mismatches: 16
Query Match: 34.67% Indels: 4
DB: 2 Gaps: 3
AF334735 (1-954) x I46506 (1-146)
QY 136 ATGTCGATTCATTTCTCCACACCCACTACCGAATTCACAGGATTTGGGAATCTTCTT 195
Db 1 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeu 20
QY 196 GAAGGGCTGACCGGAGATTCTGAGAGCAACCGGCAATATATACCGCTTTTCAGCA 255
Db 21 GluGlyLeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAlaAla 40
QY 256 GCCTATTTTGTAGAGCCTTCTAGAGAAAAGAGAGAAAACCAACTTTGTATCCAGCAGAAATGG 315
Db 41 AlaTyrPheGluSerLeuLeuGluLysArgGluLysThrAsnPheAspProAlaGluTrp 60
QY 316 GGGAGTAAGGTAGAGACCGCTTCTATTAACAATCATGCAATTCGAGGAGCAAGAACCACT 375
Db 61 GlySerLysValGluAspArgPheTyrAsnAsnHisAlaPheGluGluGlnPro 80
QY 376 GAGAAAGTGTCTTAAACAGAGTCTCAGATATCTGGGAAAGGAGGAGAGACATCA 435
Db 81 GluLysSerAspProLysGlnGluSerGlnIleSerGlyLysGluGluGluThrSer 100
QY 436 GTCACCATCTTAGACTCTTCTGAGGAAGATAGGAAAAGAGAGCTTGTCTGTCAAA 495
Db 101 ValThrIleLeuAspSerSerGluGluAspLysGluLysGluLysValAlaAlaValLys 120
QY 496 ATCCAAGTCTGCTCCGGGACACATAGCCAGAGAGGCGCAAGAAAATGAAAACAAAT 555
Db 121 IleGlnAlaAlaPheArgGlyHisIleAlaArgGluGluAlaLysLysMetLysThrAsn 140
QY 556 AGTCTTCAAAATGAGGAAAAGAGGAAACAG 588
Db 141 SerLeuGlnAsnGluLysGluLysGluLys 151
RESULT 3
S49527
zona pellucida binding protein Sp17 [similarity] - mouse
C:Species: Mus musculus (house mouse)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S49527
R:Kong, M.; Richardson, R.T.; O'Rand, M.G.
submitted to the EMBL Data Library, October 1994
A:Description: Sequence and localization of the mouse sperm protein, Sp17.
A:Reference number: S49527
A:Accession: S49527
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-149 <KON>
A:Cross-references: UNIPROT:Q62252; EMBL:Z46299; NID:g561527; PIDN:CAA86455.1; PID:g561527
C:Superfamily: sperm surface protein Sp17
Alignment Scores:
Pred. No.: 1,266-43 Length: 149
Score: 558.00 Matches: 107
Percent Similarity: 86.49% Conservative: 21
Best Local Similarity: 72.30% Mismatches: 18
Query Match: 34.00% Indels: 2
DB: 2 Gaps: 1
AF334735 (1-954) x S49527 (1-149)
QY 136 ATGTCGATTCATTTCTCCACACCCACTACCGAATTCACAGGATTTGGGAATCTTCTT 195
Db 1 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeu 20
QY 196 GAAGGGCTGACCGGAGATTCTGAGAGCAACCGGCAATATATACCGCTTTTCAGCA 255
Db 21 GluGlyLeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAlaAla 40
QY 256 GCCTATTTTGTAGAGCCTTCTAGAGAAAAGAGAGAAAACCAACTTTGTATCCAGCAGAAATGG 315
Db 41 AlaTyrPheGluSerLeuLeuGluLysArgGluLysThrAsnPheAspProAlaGluTrp 60
QY 316 GGGAGTAAGGTAGAGACCGCTTCTATTAACAATCATGCAATTCGAGGAGCAAGAACCACT 375
Db 61 GlySerLysValGluAspArgPheTyrAsnAsnHisAlaPheGluGluGlnPro 80
QY 376 GAGAAAGTGTCTTAAACAGAGTCTCAGATATCTGGGAAAGGAGGAGAGACATCA 435
Db 81 GluLysSerAspProLysGlnGluSerGlnIleSerGlyLysGluGluGluThrSer 100
QY 436 GTCACCATCTTAGACTCTTCTGAGGAAGATAGGAAAAGAGAGCTTGTCTGTCAAA 495
Db 101 ValThrIleLeuAspSerSerGluGluAspLysGluLysGluLysValAlaAlaValLys 120
QY 496 ATCCAAGTCTGCTCCGGGACACATAGCCAGAGAGGCGCAAGAAAATGAAAACAAAT 555
Db 121 IleGlnAlaAlaPheArgGlyHisIleAlaArgGluGluAlaLysLysMetLysThrAsn 140
QY 556 AGTCTTCAAAATGAGGAAAAGAGGAAACAG 588
Db 141 SerLeuGlnAsnGluLysGluLysGluLys 151

Db 21 GluGlyLeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAlaAla 40
QY 256 GCCTATTTTGTAGAGCCTTCTAGAGAAAAGAGAGAAAACCAACTTTGTATCCAGCAGAAATGG 315
Db 41 AlaTyrPheGluSerLeuLeuGluLysArgGluLysThrAsnPheAspProAlaGluTrp 60
QY 316 GGGAGTAAGGTAGAGACCGCTTCTATTAACAATCATGCAATTCGAGGAGCAAGAACCACT 375
Db 61 GlySerLysValGluAspArgPheTyrAsnAsnHisAlaPheGluGluGlnPro 79
QY 376 GAGAAAGTGTCTTAAACAGAGTCTCAGATATCTGGGAAAGGAGGAGAGACATCA 435
Db 81 GluLysSerAspProLysGlnGluSerGlnIleSerGlyLysGluGluGluThrPro 97
QY 436 GTCACCATCTTAGACTCTTCTGAGGAAGATAGGAAAAGAGAGCTTGTCTGTCAAA 495
Db 98 ValThrIleLeuAspSerSerGluGluAspLysGluLysGluLysMetAlaAlaLeuLys 116
QY 496 ATCCAAGTCTGCTCCGGGACACATAGCCAGAGAGGCGCAAGAAAATGAAAACAAAT 555
Db 117 IleGlnAlaAlaPheArgGlyHisLeuAlaArgGluLysValLysLysIleArgThrAsn 136
QY 556 AGTCTTCAAAATGAGGAAAAGAGGAAAC 585
Db 137 LysAlaGluGluGluThrGluGluAsnAsn 146
RESULT 3
S49527
zona pellucida binding protein Sp17 [similarity] - mouse
C:Species: Mus musculus (house mouse)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S49527
R:Kong, M.; Richardson, R.T.; O'Rand, M.G.
submitted to the EMBL Data Library, October 1994
A:Description: Sequence and localization of the mouse sperm protein, Sp17.
A:Reference number: S49527
A:Accession: S49527
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-149 <KON>
A:Cross-references: UNIPROT:Q62252; EMBL:Z46299; NID:g561527; PIDN:CAA86455.1; PID:g561527
C:Superfamily: sperm surface protein Sp17
Alignment Scores:
Pred. No.: 1,266-43 Length: 149
Score: 558.00 Matches: 107
Percent Similarity: 86.49% Conservative: 21
Best Local Similarity: 72.30% Mismatches: 18
Query Match: 34.00% Indels: 2
DB: 2 Gaps: 1
AF334735 (1-954) x S49527 (1-149)
QY 136 ATGTCGATTCATTTCTCCACACCCACTACCGAATTCACAGGATTTGGGAATCTTCTT 195
Db 1 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeu 20
QY 196 GAAGGGCTGACCGGAGATTCTGAGAGCAACCGGCAATATATACCGCTTTTCAGCA 255
Db 21 GluGlyLeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAlaAla 40
QY 256 GCCTATTTTGTAGAGCCTTCTAGAGAAAAGAGAGAAAACCAACTTTGTATCCAGCAGAAATGG 315
Db 41 AlaTyrPheGluSerLeuLeuGluLysArgGluLysThrAsnPheAspProAlaGluTrp 60
QY 316 GGGAGTAAGGTAGAGACCGCTTCTATTAACAATCATGCAATTCGAGGAGCAAGAACCACT 375
Db 61 GlySerLysValGluAspArgPheTyrAsnAsnHisAlaPheGluGluGlnPro 80
QY 376 GAGAAAGTGTCTTAAACAGAGTCTCAGATATCTGGGAAAGGAGGAGAGACATCA 435
Db 81 GluLysSerAspProLysGlnGluSerGlnIleSerGlyLysGluGluGluThrProVal 98

QY 436 GTCACCATTTAGACTCTTCTGAGGAGATAGAGAAAAGAGAGGTTGCTCTCTCAAA 495
 Db :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 QY 99 ThrProPheGluSerThrGluGluArgGluGluGluGluGluGluGluGluGluGlu 118
 Db :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 QY 496 ATCCAGCTGCTTCGGGGACACATAGCCAGAGAGGCGCAAGAAATGAAACAAAT 555
 Db :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 QY 119 IleGlnSerLeuPheArgGlyHisValAlaArgGluGluValLysMetLysSerAsp 138
 Db :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 QY 556 AGCTCTCAAAATGAGGAAAGAG 579
 Db :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 QY 139 LysAsnGluAsnLeuLysGluGlu 146
 Db :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 RESULT 4
 T29159
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T29159
 R:Minx, P.; Wohldmann, P.
 A:Description: The sequence of C. elegans cosmid F39H12.
 A:Reference number: Z20580
 A:Accession: T29159
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-189 <MIN>
 A:Cross-references: UNIPROT:Q20199; EMBL:U42847; PIDN:AAA83605.1; CESP:F39H12.3
 C:Genetics:
 A:Gene: CESP:F39H12.3
 A:Introns: 49/1; 70/3; 96/3; 126/2
 C:Superfamily: sperm surface protein Sp17
 Alignment Scores:
 Pred. No.: 0.00321 Length: 189
 Score: 134.00 Matches: 47
 Percent Similarity: 39.2% Conservative: 26
 Best Local Similarity: 25.2% Mismatches: 44
 Query Match: 8.1% Indels: 69
 DB: 2 Gaps: 7
 AF334735 (1-954) x T29159 (1-189)
 QY 169 APTCCACAGAGATTTGGGAATCTTCTGAAGGCTGACACGCGAGATCTTGAGAGCA 228
 Db :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 QY 229 CCGGACATATACAGCTTTTCAGCAGCGCTTATTTGAGCGCTTCTAGAG---AAAGA 285
 Db :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 QY 29 ProSerAspValAlaGluPheGlyHisMetPheAspGluTyrLeuLysHisArgArg 48
 Db :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 QY 286 GAGAAACCAACTTT-----CATCCAGCAGAAATGGGGGAGTAAGTAGAAGACCGCTTC 339
 Db :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 QY 49 GluAsnArgAsnIleLeuLysAspProAla----- 58
 Db :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 QY 340 TATAACAATCATGCATTCGAGGAGCAAGAACCCCTGAGAAAGTGCCTTAAACAAGAA 399
 Db :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 QY 59 -----AlaTyrGlu-----ValPheArgAlaAspLeuGlnLysLys 70
 Db :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 QY 400 GAGTCTCAGATATCTGGGAAGGAGGAAGACATCAGTCACCATCTTAGACTCTTCTGAG 459
 Db :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 QY 71 PheAlaGluVal-----Glu 75
 Db :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 QY 460 GAAGATAAGGAAAGAGAGCTTCTGCTGTGTCAAAATCCAAAGTGCCTTCGGGGACAC 519
 Db :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 QY 76 ArgProAlaSerProMetAspThrAlaAlaThrLysIleGlnAlaAlaPheLysGlyHis 95
 Db :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 QY 520 ATAGCCAGAGAGGAGCAAGAAATG-----AAA 549
 Db :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 QY 96 LeuValArgAlaHisProGluLysTyrGlyMetSerThrArgThrSerSerSerGluLys 115
 Db :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 QY 550 ACAATAGTCTTCAAAATGAGGAAAGAGGAAA----- 583
 Db :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 116 LeuAspSerAlaAsnAsn-LysLysAspGlnLysArgHisSerValGlyGlyTyrThrI 135
 QY 584 -----ACAACTGAGGACACTGGTTT 603
 Db 135 eAspValAspThrProGluAspArgAlaAlaThrLysIleGlnSerGluLeuArgGlyP 155
 QY 604 TACCTCCAGGAACAT 619
 Db 155 eLeuThrArgLysHis 160
 RESULT 5
 ORK2UR
 protein kinase (EC 2.7.1.37), cAMP-dependent, type II-alpha regulatory chain - human
 C:Species: Homo sapiens (man)
 C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004
 C:Accession: S03885
 R:Oeyen, O.; Myklebust, F.; Scott, J.D.; Hansson, V.; Johnsen, T.
 FEBS Lett. 246, 57-64, 1989
 A:Title: Human testis cDNA for the regulatory subunit RII-alpha of cAMP-dependent protei
 A:Reference number: S03885; MUID:89211413; PMID:2540040
 A:Accession: S03885
 A:Molecule type: mRNA
 A:Residues: 1-404 <OYE>
 A:Cross-references: UNIPROT:P13861; EMBL:X14968; NID:G29647; PIDN:CAA33094.1; PID:G29648
 C:Comment: The inactive form of the enzyme is composed of two regulatory chains and two
 our cAMP molecules.
 C:Comment: Four types of regulatory chains are found: I-alpha, I-beta, II-alpha, and II-
 C:Comment: Type II regulatory chains are phosphorylated by the activated catalytic chain
 C:Comment: Type II regulatory chains mediate membrane association by binding to anchorin
 merization.
 C:Genetics:
 A:Gene: GDB:PRKAR2A
 A:Cross-references: GDB:120314; OMIM:176910
 A:Map position: 7pter-7p22
 C:Superfamily: cAMP-dependent protein kinase regulatory chain; cAMP receptor protein cys
 F:2-404/Product: protein kinase, cAMP-dependent, type II-alpha regulatory chain #status
 F:2-138/Domain: protein interaction <DIM>
 F:139-260/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CAP1>
 F:261-393/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CAP2>
 F:2/Modified site: acetylated amino end (Ser) (in mature form) #status predicted
 F:99/Binding site: phosphate (Ser) (covalent) (by autophosphorylation) #status predicted
 F:208,217/Binding site: cAMP (Glu, Arg) #status predicted
 F:338,347/Binding site: cAMP (Glu, Arg) #status predicted
 Alignment Scores:
 Pred. No.: 0.00972 Length: 404
 Score: 118.50 Matches: 38
 Percent Similarity: 41.26% Conservative: 21
 Best Local Similarity: 26.57% Mismatches: 53
 Query Match: 7.22% Indels: 31
 DB: 1 Gaps: 4
 AF334735 (1-954) x ORK2UR (1-404)
 QY 157 ACCCACTACCGAATTCACAAAGGATTTGGGAATCTTTGAGGCTGCACGCGAGATT 216
 Db 2 SerHisIleGlnIleProGlyLeuThrGluLeuLeuGlnGlyTyrThrValGluVal 21
 QY 217 CTGAGAGAGCAACCGGACAAATATACCACTTTTGAGCAGCGCTATTTTGAGAGCGCTTCTA 276
 Db 22 LeuArgGlnGlnProProAspLeuValGluPheAlaValGluTyrPheThrArgLeuArg 41
 QY 277 GAGAAAGAGAGAGAAACCAACTTTTCATCCAGCAAGAACCACTCGAGGAGTAAGTAGAAGACCGC 336
 Db 42 GluAlaArgAlaProAlaSerValLeuProAla-----AlaThrProArg 56
 QY 337 TTCTATAACAATCATGCATTCGAGGAGCAAGAACCACTCGAGGAGTAAGTAGAAGACCGC 396
 Db 57 GlnSerLeuGlyHisProProGluProGlyProAspArgValAlaAspAlaLysGly 76
 QY 397 GAAGAGTCTCAGATATCTGGGAAGGAGGAGACATCAGTCACCATCTTAGACTCTTCT 456
 Db :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

C:Superfamily: cAMP-dependent protein kinase regulatory chain; cAMP receptor protein cyd
 C:Keywords: autophosphorylation; cAMP binding; duplication; heterotrimer; homodimer; P
 F:2-403/Product: protein kinase, cAMP-dependent, type II regulatory chain #status predic
 F:2-155/Domain: protein interaction <DM>

F:156-278/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CAP1>
 F:179-403/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CAP2>
 F:117/Binding site: phosphate (Ser) (covalent) (by autophosphorylation) #status predicted
 F:226.235/Binding site: cAMP (Glu, Arg) #status predicted
 F:349.358/Binding site: cAMP (Glu, Arg) #status predicted

Alignment Scores:
 Pred. No.: 0.0919 Length: 403
 Score: 108.00 Matches: 32
 Percent Similarity: 40.62% Conservative: 20
 Best Local Similarity: 25.00% Mismatches: 68
 Query Match: 6.58% Indels: 8
 DB: 1 Gaps: 1

AF334735 (1-954) x A43435 (1-403)

QY 163 TACCAATTCACACAGGATTGGGAATCTTCTTGAAGGCTGCACACGCGAGATTCGTGAGA 222

Db 4 TyrThrIleProSerGluLeuProProlIleLeuLysAspLeuSerArgGluValLeuArg 23

QY 223 GAGCAACCGGACATATACCACTTTTGCAGCAGCTATTTTGAGAGCCTTCTAGAGAAA 282

Db 24 HisGlnProAlaAspLeuValGlnPheCysHisAspTyrPheAlaLysLeuLeuAlaGln 43

QY 283 AGAGAGAAAACCACTTTGATCCACAGAGTAAGGAGTAAAGTAAAGAGCGCTTCTAT 342

Db 44 GlnArgLysValLeuMetAspSerAlaAspProAlaThrLysAlaThrIleAlaSerThr 63

QY 343 AACATCATGCTTCAGAGGACAGAACCACTCGAAGAGTAAAGTAAAGTAAAGAGAG 402

Db 64 AlaGlyProAlaValAlaAspAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 83

QY 403 TCTCAGATATCTGGGAAGGAGGAGAGACATCAGTCACCATCTTATAGCTTCTTCAGGAA 462

Db 84 AspAspGlyPheGlyThrGlu-----AspAspAspAsp 95

QY 463 GATAGGAAAAGAGAGGTTGCTCTGTCAAAATCCAAGTGCCTTCGGGGGACACATA 522

Db 96 AspAspAspGluAspGluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 115

QY 523 GCCAGAGGAGGCGAAGAAATG 546

Db 116 ThrSerValSerAlaGluSerMet 123

RESULT 11
 S37932
 hypothetical protein YKL105c - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein YKL459
 C:Species: Saccharomyces cerevisiae
 C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
 A:Accession: S37932; S39099
 R:Cheret, G.; Fukuhara, H.; Bolotin-Fukuhara, M.; Daignan-Fornier, B.; Pallier, C.; Puzo
 submitted to the Protein Sequence Database, March 1994

A:Reference number: S37920
 A:Accession: S37932
 A:Molecule type: DNA
 A:Residues: 1-1132 <CHR>

A:Cross-references: UNIPROT:P34250; EMBL:Z28105; NID:G486176; PIDN:CAA81945.1; PID:G4861
 A:Experimental source: strain S288C
 R:Cheret, G.; Pallier, C.; Valens, M.; Daignan-Fornier, B.; Fukuhara, H.; Bolotin-Fukuha
 Yeast 9, 1259-1265, 1993

A:Title: The DNA sequence analysis of the HAP4-LAP4 region on chromosome XI of Saccharom
 A:Reference number: S39094; MUID:94152173; PMID:8109175
 A:Accession: S39099

A>Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1132 <CHE>

A:Cross-references: EMBL:X71133; NID:G431205; PIDN:CAA50452.1; PID:G431211
 A:Experimental source: strain S288C

C:Genetics:
 A:Cross-references: SGD:S0001588
 A:Map position: 11L

Alignment Scores:

Pred. No.: 0.104 Length: 1132
 Score: 108.00 Matches: 48
 Percent Similarity: 42.53% Conservative: 46
 Best Local Similarity: 21.72% Mismatches: 83
 Query Match: 6.58% Indels: 44
 DB: 2 Gaps: 8

AF334735 (1-954) x S37932 (1-1132)

QY 67 AAACAACCGGAACCGCGGCACACGCTCGAGAGAGAAAGGAGTTCCATAGCAGTTCTT 126

Db 202 LysThrThrLysLysAlaGlySerGlnGluLysLysSerGlu----- 215

QY 127 ACCAAGAAGATGTCGATTCCATTCTCCAACACCCACTACCGAATTCACAAAGGATTTGGG 186

Db 216 -----SerGlyGlySerLysSerLysSerLysHis-----Gly 226

QY 187 AATCTTCTTGAAGGCTGCACACGCGAGATTCTGAGAGCAACCGGACAAATATACCACT 246

Db 227 LysValLeuSerLysMetHisThrThrSerLeuLysGlnArgHisAsnValProLeu 246

QY 247 TTTGACGAGCCTATTTTGAGAGCCTTCTAGAGAAAGAGAGAAACCACTTTGATCCA 306

Db 247 IleProThrThrMetAsnGluGluThrGluGlnGluLeuGlnAspLeuHisAspPro 266

QY 307 GCAGAATGGGGAGTAGGTAGAGACCGCTTCTATAACAATCATGCTTCGAGGAGCAA 366

Db 267 LeuGluPheLysProMetIleIleSerAsp-----AspGluAsnSerPheIleGluPro 284

QY 367 GAACACCTCGAGAAAAGTGATCTCTAAACAAGAGAGTCTCAGATATCTGGGAAGAGGAA 426

Db 285 SerValLeuAspArgSerIleProLysLysAspLysSerGlyLeuSerGlyArgGluLys 304

QY 427 GAG-----ACATCA 435

Db 305 LysGluGluIleGluThrLeuLeuLysGluValHisSerLeuGluGluLysIleSerAsn 324

QY 436 GTCCACCTCTTAGACTCTTCTGAGAGAGATGAAGGAAAAGAA-----GAGTTGCT 486

Db 325 IleGluIleAlaLysLeuAsnGluGluGluArgGluGlnSerLeuIleLeuGluLeuArg 344

QY 487 GCTGTCAAAATCCAAGTGCCTTCGGGGACACATACGACAGAGAGGAGGAGGAGGAGGAG 546

Db 345 LysValLysLeuAspGluGluArgMetGluLeuLysArgGluLeuAsnIleVal 364

QY 547 AAACAATAAGT---CTTCAAAATGAGGAAAA-----GAGGAAAAACAGTGAGGA 594

Db 365 LysGluAsnAlaAspLeuGluAlaGlnGluLeuLysLeuIleGluSerLysArgLysGln 384

QY 595 CACTGGTTTACCTCCAGGAACA-----TGAAAAATAATCCAATCATCAACTCATCACT 645

Db 385 HisPheHisLysGlyGlnGluValAlaSerGluValLysSerIleThrIleArgGlnPro 404

QY 646 TCT 648

Db 405 Thr 405

RESULT 12

E70474

translation initiation factor IF-2 - Aquifex aeolicus

C:Species: Aquifex aeolicus

C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004

C:Accession: E70474

R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov

V.

Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A:Reference number: A70300; MUID:98196666; PMID:9537320

A:Accession: E70474
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-805 <AQF>
A:Cross-references: UNIPROT:O67825; GB:A8000769; NID:g2984262; PIDN:AAC07794.1; PID:g2984262
A:Experimental source: strain VF5
C:Genetics:
A:Gene: infB
C:Superfamily: Translation initiation factor IF-2; translation elongation factor Tu homolog
C:Keywords: GTP binding; nucleotide binding; P-loop
F:309-419/Domain: translation elongation factor Tu homology <ETU>
F:315-322/Region: nucleotide-binding motif A (P-loop)
F:416-419/Region: GTP-binding NXXD motif
F:452-454/Region: GTP-binding SAK/L motif
F:321,322,342,416,417,419,452/Binding site: Mg-GTP (Lys, Thr, Thr, Asn, Lys, Asp, Ser) #

Alignment Scores:			
Pred. No.:	0.111	Length:	805
Score:	107.50	Matches:	53
Percent Similarity:	42.49%	Conservative:	29
Best Local Similarity:	27.46%	Mismatches:	74
Query Match:	6.55%	Indels:	37
DB:	2	Gaps:	8
AF334735 (1-954) x E70474 (1-805)			
Qy	73	ACCGGAACCGCGGCACACGACTCGGAGAGAAAGAGGTTCCATAGGAGTTCATTACCAAG	132
Db	4	ThrlYsThrLysArgValSerAspValAlaLysGluLeu-----GlyValLysSerLys	21
Qy	133	AAGATGTCGATTCCATTCTCAACACCCACTTACCGAATCCACAAGA-----	180
Db	22	GluLe-----IleGluPheLeuAsnGluTyrTrpProArgProAspGlyLysProTrpLys	40
Qy	181	-----TTTGGGAATCTT	192
Db	41	AlaSerHisGlyLeuAspGluGlnAlaLeuGluMetIleTyrAspAlaPheGlyIleLys	60
Qy	193	CTTGAAGGGCTGACACCGGAGATTCTCAGAGAGCAACCGCAATATACACGCTTTTGCA	252
Db	61	GluGluGluGluLysGluGluValIleThrGluGln-----AlaGlnAla	75
Qy	253	GCAGCCTATTTGAGAGCCTTCTAGAGAAAGAGAGAAAAACCAACTTTGTATCCAGCAGAA	312
Db	76	ProAlaGluValGlu-----GluLysGluGluGluLysGluGluValIle	92
Qy	313	TGGGGGAGTAGAGAGCCGCTTCTATACAAATCATGCTATTCGAGGAGCAGAAGCA	372
Db	93	ValGluGluValValGluGluLys-----LysProGluValIleValGluGluIleGluGlu	111
Qy	373	CCTGAGAAAAGTGATCTCTAAACAAGAGAGTCTCAGATATCTGGGAAGGAGGAAGACA	432
Db	112	LysLysGluGluGluLysLysGluGluLysGluGluLysProLysLysSerValGluGluLeu	131
Qy	433	TCAGTCAACCATCTTAGACTCTCTCGAGGAAGATAAGGAAAAAGAGAGGTTGCTGCTGTC	492
Db	132	IleLysGluIleLeuGluLysLysGluLysGluLysGluLysLysLysValGluLysGlu	151
Qy	493	AAATCCAAAGCTGCTCTTCGG-----GGACACATAGCCAGGAGGAGGCAAGAAAAATCAAA	549
Db	152	ArgLysGluGluLysValArgValValGluValLysLysGluGluArgLysGluGluLys	171
Qy	550	ACAAATAGTCTTCAAAATGAGGAAAAAGAGCGAAAAACAAAG	588
Db	172	-----LysGluGluLysLysGluGluLysLys	180

RESULT 13
S78475
mannosylphosphorylation protein MN4 - yeast (*Saccharomyces cerevisiae*)
N;Alternate names: protein YKL200c; protein YKL201c
C;Species: *Saccharomyces cerevisiae*
C;Date: 11-Mar-1998 #sequence revision 17-Apr-1998 #text_change 09-Jul-2004
C;Accession: S78475; S38037; S38038

R:Odani, T.; Shimma, Y.; Yoshifumi, J.
submitted to the EMBL Data Library, January 1996
A:Description: Cloning and Analysis of the MNNA Gene Required for Phosphorylation of N-linked oligosaccharides
A:Reference number: S78475
A:Accession: S78475
A:Molecule type: DNA
A:Residues: 1-1178 <DA>
A:Cross-references: UNIPROT:P36044; EMBL:D83006; NID:G1752735; PID:d1012343; PID:g17527
A:Note: This is a revision to the sequence from reference S38024
R:Malá e Silva, A.; Bossier, P.; Vilela, C.; Fernandes, L.; Soares, H.; Guerreiro, P.;
submitted to the Protein Sequence Database, March 1994
A:Reference number: S38024
A:Accession: S38037
A:Molecule type: DNA
A:Residues: 121-249
A:Note: This sequence has been revised in reference S78475
A:Cross-references: EMBL:D28200; NID:G486355; PID:g486356
A:Experimental source: strain S288C
A:Note: this sequence has been revised in reference S78475
A:Note: this was assumed to be protein YKL200C
A:Accession: S38038
A:Molecule type: DNA
A:Residues: 640-1178 <MA>
A:Cross-references: EMBL:D28201; NID:G486357; PID:g486359; MIPS:YKL201C
A:Experimental source: strain S288C
A:Note: this sequence has been revised in reference S78475
A:Note: this was assumed to be the complete sequence of protein YKL201C
C:Genetics:
A:Gene: SGD:MNNA4; MNNA
A:Cross-references: MIPS:YKL201C; SGD:S0001684
A:Map position: 11L
C:Function:
A:Description: required for phosphorylation of N-linked oligosaccharides
C:Keywords: transmembrane protein
F:28-44/Domain: transmembrane #status predicted <TMM>

```
AF334735 (1-954) x S78475 (1-1178)

Qy      133 AAGATGTCGATTCCCATTTCTCCCAACACCCTACTACCGAATTCACAAGGATTTGGG---AAT   189
||| : : : : : ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      981 LysSerAsnLeuAsnPheserSerAsnSerPheLeuAspLeuAsnSerTyrGlyLeuAsp   1000
|||| : : : : : ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      190 CTTCTTGGAAGGCTGCACAGCGCAGAGATTCCTCGAGAGAG-----CAACCCGAC   234
||| : : : : : ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1001 LeuPheAlaProThrLeuSerAspValAsnArgLysGlyIleGlnMetPheAspLysAsp   1020
|||| : : : : : ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      235 AATATACCAGCTTTTGCAGCAGCCTATTTTTGAGAGCCTTCTTAGAG-----AAAAGA   285
||| : : : : : ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1021 ProIleIleValTyrGluaspTyrAlaIleLysLeuLeuGluGluArgLysArgArg   1040
|||| : : : : : ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      286 GAGAAAAACAACTTTGTATCCACGAGAAATGGGGAGTAAGCTAGAAGACCGCTTCTATAAC   345
||| : : : : : ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1041 GluLysLysLysLysGluGluGluGluLysLysLysGluGluGluLysLysLys   1060
|||| : : : : : ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      346 AATCATGCTTCGAGGAGCAGAACCACTCGAGAAAAGTGCATCTCTAAACAAGAGAGTCT   405
|||| : : : : : ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1061 LysGluGluGluGluLysLysLysLysGluGluGluLysLysLysLysGluGlu   1080
|||| : : : : : ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      406 CAGATATCTGGGAACGAGGAAGACAGATCAGTCACCATCTTAGACTCTTCTGAGGAGAT   465
||| : : : : : ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1081 GluLysLysLysLysGluGluGluGluLysLysLys-----GlnGluGluGluGluLys   1098
|||| : : : : : ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      466 AAGNAAAAAGAGAGGTGCTGCTGCTCAANAATCCAAGCTGCCCTCCGGGGACACATAGCC   525
|||| : : : : : ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1099 LysLysLysGluGluGluGluLysLysLys-----LysLysLysLysLysLysLysLys   1108
|||| : : : : : ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```


A:Residues: 1-699 <W12>
A:Cross-references: EMBL:AL021448; PIDN:CAA16278.1; GSPDB:GN00023; CBSP:Y2H9A.4
A:Experimental source: clone Y2H9A
C:Genetics:
A:Gene: CBSP:Y2H9A.4
A:Map position: 5
A:Introns: 35/3; 65/2; 112/1; 146/3; 184/3; 287/2; 325/1; 457/2; 500/3; 562/3; 577/3; 61

Alignment Scores:
Pred. No.: 0.232 Length: 699
Score: 104.00 Matches: 58
Percent Similarity: 43.48% Conservatives: 32
Best Local Similarity: 28.02% Mismatches: 69
Query Match: 6.34% Indels: 48
DB: 2 Gaps: 11

AF334735 (1-954) x T21379 (1-699)

```
QY 38 AAAGAAAAACGGTTACC---CAGCAACTAGAAAAACACCGAACCCGGCCGACACGCT 94
||| ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 354 LysPheLysThrCysSerAspGluLysLeuAspLys-SerArgArgSerLeuGlnI 373
QY 95 CGGAGAAAGAGGCTTCATAGGAGCTTCATACCAAGAGATGCGATTCTCTCCCA 154
: ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 373 eGlyGluLysSerGluThrAspGluAsnGluLeuLysGlyValaspGln-----As 391
QY 155 ACACCCACTACCGAATTCACAGGATTGGG-----AATCTTCTGAAGGCTGACAC 208
: : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 391 phenGlnLeuProAsnThrSerArg-PheSerLeuGlnAsnGluLeuAspSerLeuGluA 411
QY 209 GCGAGATCTGAGAGCAACCGGACAATATACCAAGCTTTTCACGACGCTATTTTGAGA 268
||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 411 rgThrLeuGlnLys-----AsnIleLysValaspValGlyAla-----G 424
QY 269 GCCTCTTAGAGAAA-----AGAGAGAAAACCACTTT-----GATCCAG 307
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 424 InLeuLeuGluLysTySerAsnGluLeuThrAsnPheValAlaAsnIleThrIlePro- 443
QY 308 CAGAATGGGGAGTAGGAGAGAGCGCTCTTATACCAATCATGCTGAGGAGGACGAG 367
||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 444 -----LysGlnLysGluAspIlePheArgGluThrGlu-----GluArgAspA 458
QY 368 AACCACTTGAGAAAAGTGATCTCTAAACAGAGAGTCTCAGATATCTGGGAAGGAGGAG 427
||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 458 snProSerLysSerAspSerGluAsnGluAspAsnAspPheAsnGluLysL 478
QY 428 AGACATCAGTCACCTCTAGACTCTCTGAGAGAGATAGGAAAAGAGAGGTTGCTG 487
||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 478 euLysGlnGlnGlnLysAsnGluGluGluLysLysLysArgGlnGluArgGluG 498
QY 488 CTGTCAAAATCCAAGCTGCTTCGCGGGACACATACGAGAGAGGAGGCAAGAAATGA 547
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 498 luArgLysArgGln-----AlaGlnGluLysAsnLeuL 511
QY 548 AAACAATAAGT-----CTTCAAA 565
||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 511 ysSerGluSerArgAlaArgLeuSerAlaPheMetThrCysIleArgLeuThrIleArgP 531
QY 566 ATGAGGAAAAGAGGAA 582
||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 531 heGluGlnLysGluGlu 536
```

RESULT 17

152638
Neuronal growth-related protein GAP-43 - human
N:Alternate names: B-50; growth-associated protein GAP-43; neuromodulin
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: 152638; S06437; JH0179
R:Nielsen, H.B.; De Groen, P.C.; Eggen, B.J.; Schrama, L.H.; Gispén, W.H.; Schotman, F.
Brain Res. Mol. Brain Res. 19, 293-302, 1993
A:Title: Structure of the human gene for the neural phosphoprotein B-50 (GAP-43).
A:Reference number: 152638; MUID:94049025; PMID:8231732

A:Accession: 152638
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-238 <RES>
A:Cross-references: UNIPROT:P17677; GB:S66541; NID:9440922; PIDN:AAB28649.1; PID:9440922
R:Kosik, K.S.; Orecchio, L.D.; Bruns, G.A.F.; Benowitz, L.I.; MacDonald, G.P.; Cox, D.R.
Neuron 1, 127-132, 1988
A:Title: Human GAP-43: Its deduced amino acid sequence and chromosomal localization in
A:Reference number: S06437; MUID:90166498; PMID:3272162

A:Accession: S06437
A:Molecule type: mRNA
A:Residues: 1-238 <KOS>
A:Cross-references: EMBL:M25667; NID:gl82969; PIDN:AA52516.1; PID:gl82970
R:Ng, S.C.; de la Monte, S.M.; Conboy, G.L.; Karns, L.R.; Fishman, M.C.
Neuron 1, 133-139, 1988
A:Title: Cloning of human GAP-43: growth association and ischemic resurgence.
A:Reference number: JH0179; MUID:90166499; PMID:3272163

A:Accession: JH0179
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-238 <NG>
C:Genetics:
A:Gene: GDB:GAP43
A:Cross-references: GDB:119972; OMIM:162060
A:Map position: 3q13.1-3q13.2
A:Introns: 10/3; 210/1
C:Superfamily: histone H1
C:Keywords: lipoprotein; phosphoprotein; thiolester bond
F:3,4/Binding site: palmitate (Cys) (covalent) #status predicted

Alignment Scores:
Pred. No.: 0.251 Length: 238
Score: 103.00 Matches: 29
Percent Similarity: 48.75% Conservatives: 10
Best Local Similarity: 26.25% Mismatches: 26
Query Match: 6.28% Indels: 15
DB: 2 Gaps: 1

AF334735 (1-954) x 152638 (1-238)

```
QY 351 TGCATTCGAGGAGCAAGCAACCACTGAGAAAGTGATCTTAACCAAGAGTCTCAGAT 410
||| ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 4 CysMetArg-gThrLysGlnVal-GluLysAsnAspAspGlnLysIleGluGlnAs 23
QY 411 ATCTCGGAGGAGGAGAGAGACATCAGTCACCATCTTCTGAGGAAGATAAGCA 470
||| ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 23 pGlyIleLysProGlu-----As 29
QY 471 AAAAGAGAGGTGTCTGTCAAAATCAAGCTCCCTCCGGGACACATAGCCAGAGA 530
||| ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 29 pLysAlaHisLysAlaIleThrLysIleGlnAlaSerPheArgGlyHisIleThrArgly 49
QY 531 GGAGCAAGAAATGAAACAAATAGTCTTCAAAATGAGGAAAGGAAAGCAAG 588
||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 49 sLysLysLysGlyLysLysAspValGlnAlaLysGluAlaGluAlaAsnLys 68
```

RESULT 18

S45705
Triachin 2 - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 10-Dec-1994 #sequence_revision 19-Apr-1996 #text_change 11-Jan-2000
A:Reference number: S45704; MUID:94298946; PMID:8026576
A:Accession: S45705
R:Peng, M.; Fan, H.; Kirley, T.L.; Caswell, A.H.; Schwartz, A.
FEBS Lett. 348, 17-20, 1994
A:Title: Structural diversity of triadin in skeletal muscle and evidence of its existen
A:Reference number: S45704; MUID:94298946; PMID:8026576
A:Accession: S45705
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-204 <PEN>
C:Superfamily: histone H1

Alignment Scores:

Alignment Scores:
Pred. No.: 0.274 Length: 204
Score: 102.50 Matches: 52
Percent Similarity: 42.8% Conservative: 32
Best Local Similarity: 26.53% Mismatches: 67
Query Match: 6.25% Indels: 45
DB: 2 Gaps: 10

AF334735 (1-954) x S45705 (1-204)

QY 43 AAAACGGTTACCCAGCACTAGAAAAACAACCGGAACCGCGGCACCGAGCTCGGAGAGA 102
|||
Db 2 LysHisGluGluProAla-----LysSerThrLysLysGluHisAlaGluileGluArg 19
|||
QY 103 AAGGAGGTTCCATAGGAGCTTACCAAGAGATGTCGATCCATTC----- 150
|||
Db 20 LysGluGluValSerAlaAlaSerThrLysLys---AlaValProAlaLysLysGluGlu 38
|||
QY 151 -----TCCAAACCCCACTACCGAATTCACACAGGATTTGGGAAT 189
|||
Db 39 LysThrThrLysThrValGluGluThrArgLysGluLysProGlyLysIleSerSer 58
|||
QY 190 CTTCTTGAAGGCTGACCGGAGATTCTGAGAGCAACCGGACAATATACCACTTTT 249
|||
Db 59 ValLeuLys-----AspLysGluLeuThrLysGluLysGluValLysValProAla--- 75
|||
QY 250 GCAGCAGCTATTGAGAGCTTCTAGAGAAAGAGAGAAACCACTTTGATCCAGCA 309
|||
Db 76 -----SerLeuLysGluLysGlySerGluThrLysLysAsp----- 87
|||
QY 310 GAATGGGGAGTAGAGACCGCTTCTATAACATCATGCTTCGAGGAGCAAGAA 369
|||
Db 88 GluLysThrSerLysProGluProGlnileLys-----LysGluGluLys 102
|||
QY 370 CCACCTGAGAAAGTGATCTTAACAAGAGAGTCTCAGATATCTGGGAAGAGGAA--- 426
|||
Db 103 ProGlyLysGluValLysProLysProGlnileLysLysGluGluLysPro 122
|||
QY 427 -----GAGACATCAGTCACCATCTAGACTTCTTGAGGAGAGTAAGAA 471
|||
Db 123 GluGlnAspIleMetLysProGluLysThrAlaLeuHisGlyLysProGluGluLysVal 142
|||
QY 472 AAAGAAGAGGTTGCTGTGTCATAACATCAAGCTGCTTCGCGGGACACATAGCCAGAG 531
|||
Db 143 LeuLysGlnValLysAlaValThrThrGluLys-----HisValLysProLys 158
|||
QY 532 GAGGCAAGAAATGAACAAATAGTCTTCAAAATGAGGAAAGAAAG 579
|||
Db 159 ProAlaLysLys-----AlaGluHisGlnGluLysGlu 169
|||

RESULT 19

neuromodulin - goldfish
N:Alternate names: B-50; F1; GAP-43; pp46
C:Species: Carassius auratus (goldfish)
C>Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C:Accession: JQ0075
P:Labate, M.E.; Skene, J.H.P.
Neuron 3, 299-310, 1989
A:Title: Selective conservation of GAP-43 structure in vertebrate evolution.
A:Reference number: JQ0075; PMID:90380372; PMID:2641999
A:Accession: JQ0075
A:Molecule type: mRNA
A:Residues: 1-213 <L>B>
A:Cross-references: UNIPROT:P17691; GB:M26250; NID:g212953; PID:g212954
A:Note: This protein contains potential sites for fatty acylation and membrane attachment
C:Comment: GAP-43 is a major growth cone membrane protein whose expression is widely con
C:Comment: Features of the protein are most highly conserved in vertebrate evolution.
C:Comment: The amino end (1-57) has a strictly conserved domain including protein-protei
C:Superfamily: histone H1
C:Keywords: calmodulin binding; lipoprotein; phosphoprotein; thiolester bond
F:3.4/Binding site: palmitate (Cys) (covalent) #status predicted
F:81/Binding site: phosphate (Ser) (covalent) #status predicted

Alignment Scores:
Pred. No.: 0.306 Length: 213
Score: 102.00 Matches: 32
Percent Similarity: 44.44% Conservative: 8
Best Local Similarity: 35.56% Mismatches: 21
Query Match: 6.22% Indels: 29
DB: 2 Gaps: 3

AF334735 (1-954) x JQ0075 (1-213)

QY 351 TGCATTTCGAGGAGCAAGAACCCCTCGAGAAAGTGATCTAAACAAGAGAGTCTCAGAT 410
|||
Db 4 Cysilearg-ArgThrLysProValGluLysAsnGluLysAlaAspGlnIleLysG1 23
|||
QY 411 ATCTGGGAAGGAGGAAGAGACATCAGTCACCATCTTAGACTCTTCTGAGGAAGATAAGA 470
|||
Db 23 nAspGlyThrLysProGlu-----G1 30
|||
QY 471 AAAAGAGAGGTTCTGCTGTCAAAATCCAAGTCCCTCCGGGGACACATAGCCAGAGA 530
|||
Db 30 uAsnAlaHisLysAlaAlaThrLysIleGlnAlaSerPheargGlyHisIleThrArg-- 49
|||
QY 531 GGAGGCCAAGAAATGAAA-----ACAAA 554
|||
Db 50 -----LysLysMetLysAspGluAspLysAspGlyGluAsnAspThrAlaProAspG1 67
|||
QY 555 TAGTCTTCAAAATGAGGAAAGAGGAA 582
|||
Db 67 uSerAlaGluThrGluGluLysGluGlu 76
|||

RESULT 20

T24545
hypothetical protein T05F1.11 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T24545
R:Burton, J.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19905
A:Accession: T24545
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1003 <W1>
A:Cross-references: UNIPROT:Q9XUQ9; EMBL:Z81586; PIDN:CAB04699.1; GSPDB:GN00019; CESP:T
A:Experimental source: clone T05F1
C:Genetics:
A:Gene: CESP:T05F1.11
A:Map position: 1
A:Introns: 25/2; 61/1; 113/2; 216/3; 314/2; 345/3; 401/2; 550/3; 614/3; 682/2

Alignment Scores:
Pred. No.: 0.414 Length: 1003
Score: 101.50 Matches: 47
Percent Similarity: 37.31% Conservative: 25
Best Local Similarity: 24.35% Mismatches: 78
Query Match: 6.19% Indels: 43
DB: 2 Gaps: 8

AF334735 (1-954) x T24545 (1-1003)

QY 280 AAAAGAGAGAAACCAACTTTGATCCAGCAGATGGGGAGTAGAGACCGCTTC 339
|||
Db 741 LysArgGluValGlnMetGluTyrGluArgTrpArgIleGluGluGluArgIleu 760
|||
QY 340 -----TATAACAATCATGCTTCGAGGAGCAAGCAACCTCGAGAAAGTGATCTCTAAA 393
|||
Db 761 GluGluIleargargLysLysGluGluLysLysLysGluGluAspGluLys 780
|||
QY 394 CAAGAAGAGTCTCAGATATCTGG-----AAGGAGGAAGAGACATCAGTC 438
|||
Db 781 ArgargLysAlaGluGluSerGluArgValargLysLeuLysGluGluArgLys 800
|||
QY 439 ACCATCTTAGACTTCTCTGAGGAAGATAAGGAAAAAGAGAGGTTGCTGCTGCTCAAAATC 498
|||

418	QY	AAGGAGGAA-----GAGACATCAGTC 438
		::: :::
573	Db	LysGluGlnArgGluLeuGluGlnLeuProGlnLeuGluGluThrLysIle 592
		::: :::
439	QY	ACCATCTTAGACTCTCTGAGGAGAT-----AAGGAAAAAGAGGTTTGCT 486
		::: :::
593	Db	ThrValLeuGlnAlaGluIleAlaAspProAlaPhePheGlnGlnAlaHisAspIleThr 612
		::: :::
487	QY	GCTGTCAAAATCCAAAGCTGCTCTCCGGGGACACATAGCCAGAGAGGAGCAAGAAAATG 546
		::: :::
613	Db	AspAlaLysLeuLysAla-----LeuAlaAspThrGluAla--GluLeu 626
		::: :::
547	QY	AAACAAATAGTCTTCAAAATAGAGAAAAGAGAAAACAG 588
		::: :::
627	Db	GluThrAlaPheLeuArgTrpGluGluLeuGluLysLys 640
		::: :::

C/Accession: E64170; T09427
C/Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.H.;
Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodok, A.; Kelley, J.M.; Weidman, J.
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
Science 259, 496-512, 1995
A/Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A./Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.
A/Reference number: A64000; MUID:95350630; PMID:7542800

A:Molecule type: DNA
A:Residues: 1-647 <TIGR>
A:A/Cross-references: UNIPROT:Q57242; GB:U32809; GB:I42023; NID:G3212217; PIDN:AAC22945.1; R:White, O.; Clayton, R.A.; Kurlavag, A.R.; Fleischmann, R.D.; Peterson, J.; Hickey, E. submitted to the EMBL Data Library, May 1998
A:Reference number: Z16667
A:Accession: T09427
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-647 <WHI>
A:A/Cross-references: EMBL:U32809; NID:G3212217; PID:G1574757
C:Genetics:
A:Gene: uup-A; H11300

Alignment Scores:		
Pred. No.:	0.436	Length: 647
Score:	101.00	Matches: 39
Percent Similarity:	44.16%	Conservative: 29
Best Local Similarity:	25.32%	Mismatches: 34
Query Match:	6.15%	Indels: 52
DB:	2	Gaps: 8

217	CTGAGAGAGCAACCGGACAATATACCACTTTTTCGACGACCCCTATTTTGTAGAGCCCTCTTA	276
509	LeuPheGluGlyLysGlyHisLeuAsnLysTyrValGlyGlyPhePheAla----	526
277	GAGAAAGAGAGAAAACCACTTTGTATCCAGCAGATGGGGGAGTAGGTAGAGAACCGC	336
527	---LysGlnGlnGlnAlaAsnPhe-----TrpAlaSerLys-----	537
337	TTCTATACCAATCATGCTATTCGAGGCAAGAACCCCTCGAGAAAAGTGTATCTTAAACAA	396
538	-----AlaValGluGlnAlaLysAlaLysLysSerGluProLeuLys	552
397	GAAGAGTCT-----CAGATATCTCGG	417
553	GluGluSerAlaValLysAsnAspArgThrSerLysProLysSerValLysLeuSerTyr	572

R; Dedman, K.; Brown, D.; Bowman, S.
submitted to the EMBL Data Library, June 1995

A;Accession: S55101

A;Cross-references: UNIPROT:Q03661; EMBL:Z498

R; Skelton, J.; Churcher, C.M.

A:Accession: S57587
A:reference number: S57587

A;Residues: 608-1658 <SKE>

A;Experimental source: strain AB97Z
C:Genetics:

pred. No.:	0.49	Length
score:	101.00	Matches

Best Local Similarity: 24.10% Mismatch

DB: 2
GapB: 2[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

104 AGGAGGTTCCATAGGCAGTTCT-----

Db 1440 Arg-GlyAsnIleProSerThrAspLeuPro

QY -----TACCAAGAGAGATGTC
126 -----

[illegible]

nb 1479 rserprogluivaltyrgrlll-lepheserAs

Qy 230 CGGACAATATACCAGCTTTTGCAGCAGCCTA

db 1499 snAspGluIleProAla-----

THE

462	Qy	TTTCCTCAGAAGAGCTAAAGATGGTGACTGATGCTCTTCTCTCTCTCCGAGATATCTGAA	403
91	Db	ArgAsnArgArg-----AlaIlePheLeuLeuSerTyrPheIleAsp	104
		:::	
		:::	
402	Qy	CTCTCTCTGTTTAGGATCACTTTTCTCAGGTGGTCTTGTCTCTCGAATGTCATGTGT	343
		:::	
		:::	
105	Db	ValLeuLeuIleGlyLeuPheLeu-----LeuIle	115
342	Qy	ATAGAAGCGGTCTTCTACTCT-----ACTCCCCCATTCCTCGATCAAAAGTTGGT	292
116	Db	AlaGluLeuValPheLeuLeuSerSerPheThrProPheIleThrTyrPheSerIlePhe	135
291	Qy	TTTCTCTCTTTTCTCTAGAGGCTCTCAAATAGGTGCTGCGAAAAGCTGGTGATATGTCT	232
		:::	
		:::	
136	Db	IleSerGluTyrLeu-----LeuLeuCysAsnSerSerTyrLeuIle	149
231	Qy	CGGTGTGCTCTCAGAAATCTCGGTGTGTCAGCCCTTCAAGAAG-----ATTCCCAATCC	178
		:::	
		:::	
150	Db	AlaIleLeuIleLys-----ArgProPheArgSerPhePheIleSerIleSer	165
177	Qy	TTGTGGAATTCCGTAGTAGTGGGTGGTGGAGAATGGAATCGACATCTTCTTGTGTAGAACTGC	118
166	Db	LeuIlePheLeuValLeuGlyLe-----GluIleThrGlyAsnMetIle	180
117	Qy	CTATGGAACCTCCTT	103
181	Db	IlePheAsnTyrLeu	185

RESULT 27
G88436 protein T04A8.13 [imported] - Caenorhabditis elegans C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: G88436 R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes A;Reference number: A75000; PMID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 283, 2104, 1999;
A;Accession: G88436
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-762 <SFO>
A;Cross-references: UNIPROT.Q22142; GB.chr_III; PIDN.CAA84732.1; PID:g3879347; GSFPDB.GNC001111
C;Genetics:
A;Gene: T04A8.13
A;Map position: 3

Alignment Scores:	
Pred. No.:	0.551
Score:	100.00
Length:	762
Matches:	33
Percent Similarity:	46.67%
Conservative:	16
Best Local Similarity:	31.43%
Mismatches:	44
Query Match:	6.09%
Indels:	12
Gaps:	3
DB:	2

AF334735 (1-954) x G88436 (1-762)

Qy	277	GAGAAAGAGAGAAACCAACTTTGATCCAGCAGAATGGGGAGTAAAGTATAGAGAACGC	336
Db	77	GlulysGluylsGluclulysGluaspGlyHisGluIuylsGluaspLys	96
Qy	337	TTCTATAACAATCATGCTTCGAGAGCAAGAACCACTGAGAAAGT---GATCTTAA	393
Db	97	LysGluaspLysGluaspGluasnAspGluIuylsGluLysSerLysAspAspLys	116
Qy	394	CAAGAGAGCTTCAGATATCTGGGAAGCAGAGACATCAGTCACCATCTTAGACTCT	453
Db	117	LysGluLysLysGluaspLysThrLysThrGluaspAsnGluGly	135
Qy	454	TTTGAGGAAGATAAGAAAGAGAGGTCTGCTGTCAAAATCCAAAGTCGCTTCGG	513

```

Db      136  LysAsnGluGluLysLysLysGluLysGluGluLysLys----- 149
      |||:::||||:|||||:::|||||
Qy      514  GGACACATAGCCAGAGGAGGACAAAGAAATGAAACAAATAGCTTTCAAAATAGGAA 573
      |||:::|||||:::|||||
Db      150  -----AspGluLysGluLysLysGluLysThrGluAspAspLysGluLysGlu 165
      |||:::|||||:::|||||
Qy      574  AAAGAGGAAACAAAG 598
      |||:::|||||
Db      166  LysGluLysThrLys 170
      |||:::|||||

RESULT 28
F84514
hypothetical protein At2g14140 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: F84514
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fuj
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umavam, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.;
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: F84514
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-783 <STO>
A/Cross-references: UNIPROT:Q9SI45; GB:AE002093; NID:g4587683; PIDN:AAD25853.1
C/Genetics:
A/Map position: 2

```

[illegible]

A; Cross-references: UNIPROT:P27476; GB:X57185; NID:g4057; PIDN:CAA40472.1; PID:g4058
R; Gamberi, C.; Contreas, G.; Romanelli, M.G.; Morandi, C.
Gene 148, 59-66, 1994

A; Title: Analysis of the yeast NarI gene and protein domain comparison between NsrI and R; Skala, J.; Nawrocki, A.; Goffeau, A.

A; Reference number: PC2395; MUID:95011659; PMID:7926838

A; Accession: PC2395

A; Molecule type: mRNA

A; Residues: 155-414 <GAM>

A; Cross-references: EMBL:X57185

A; Note: The authors translated the codon TTC for residue 188 as Arg, GGT for residue 254

R; Skala, J.; Nawrocki, A.; Goffeau, A.

A; Title: The sequence of a 27 kb segment on the right arm of chromosome VII from Saccharomyces cerevisiae

A; Reference number: S60435; MUID:96158062; PMID:8585325

A; Accession: S60450

A; Status: nucleic acid sequence not shown; translation not shown

A; Molecule type: DNA

A; Residues: 1-414 <SKA>

A; Cross-references: EMBL:X85807; NID:g1045249; PIDN:CAAS9817.1; PID:g1045265

A; Note: The nucleotide sequence was submitted to the EMBL Data Library, March 1995

R; Van Dyck, L.; Skala, J.; de Wergifosse, P.; Fumelle, B.; Talla, E.; Nawrocki, A.; Delbecq, H.

A; Title: Submitted to the Protein Sequence Database, May 1996

A; Reference number: S64428

A; Accession: S64468

A; Molecule type: DNA

A; Residues: 1-414 <VAN>

A; Cross-references: EMBL:Z72944; NID:g1323270; PIDN:CAA97173.1; PID:e243542; PID:g132327

A; Experimental source: strain S288C

R; Riéger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.

A; Note: Submitted to the Protein Sequence Database, May 1996

A; Reference number: S64071

A; Accession: S64470

A; Molecule type: DNA

A; Residues: 1-188 <RIE>

A; Cross-references: EMBL:Z72944; MIPS:YGR159C

A; Experimental source: strain S288C

C; Genetics:

A; Gene: SGD:NSR1

A; Cross-references: SGD:S0003391; MIPS:YGR159C

A; Map position: 7R

C; Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoproteins; RNA binding

C; Keywords: DNA binding; nucleus; RNA binding

F; 169-236/Domain: ribonucleoprotein repeat homology <RRM1>

F; 268-336/Domain: ribonucleoprotein repeat homology <RRM2>

Db 382 LeuLeuAnValLysMetAlaLeuAsePilleGluLeuAlaAlaTyrArg----- 397
 QY 181 TTTGGGAATCTTCTGAAGGGCTGCACACGCGAGATTCTGAGAGACAAACCGGACAAATATA 240
 Db 398 -----LysLeuLeuGluGly-----GluGluThrArgPhe 407
 QY 241 CGAGCTTTTGCACGACGCTATTTTGAGAGCTTCTAGAGAAAGA----- 285
 Db 408 SerAlaPheSerGlySerIleThrGlyProIlePheThrHisArgGlnProSerValThr 427
 QY 286 -----GAGAAACCACTTTGATCCAGCAGAAATGGGGAGTAGTA 327
 Db 428 IleAlaSerThrLysIleGlnLysThrLysIleGluProProlLys-----LeuLysVal 445
 QY 328 GAAGACCGCTTCATACATCATGCTTCGAGGAGCAAGAACCACTGAGAAAAGTGTAT 387
 Db 446 GlnHisLysPheValGluGluIleIleGluGluThrLysValGluaspGluLysSerGlu 465
 QY 388 CCTAAA-----CAAGAAGAGTCT 405
 Db 466 MetGluaspAlaLeuSerAlaIleAlaGluGluMetAlaAlaLysAlaGlnGluGlu 485
 QY 406 CAGATATCTGGGAGGAGGAGACATCAGTC----- 438
 Db 486 GlnGluGluGluLysAlaGluGluGluAlaValGluGluAlaValSerGluLysAla 505
 QY 439 -----ACCATCTTAGACTCTCTGAGAGAGATAAGGAAAAGAGAGGTTGCTGCT 489
 Db 506 AlaGluGlnAlaAlaGluGluGluLysGluGluGluAlaGluGluGluGluAla 525
 QY 490 GTCAAATCCAACTGCTTCGCGGACACATAGCCAGAGAGGAGCAAGAAATGAAA 549
 Db 526 AlaLysSeraspAlaAlaGluGluGlyGlySerLysLysGluGluIleGluGluLysGlu 545
 QY 550 ACAATAGTCTTCAAAATGAGGAAAAGAGGAAAACAG 588
 Db 546 GluGlyGluGluAlaGluGluGluGluAlaGluAlaLys 558
 RESULT 33
 T02570
 hypothetical protein At2g39320 [imported] - Arabidopsis thaliana
 N;Alternate names: hypothetical protein T16B24.4
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004
 C;Accession: T02570; H84815
 R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,
 submitted to the EMBL Data Library, August 1998
 A;Description: Arabidopsis thaliana chromosome II BAC T16B24 genomic sequence.
 A;Reference number: Z14679
 A;Accession: T02570
 A;Status: translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-189 <R0U>
 A;Cross-references: UNIPROT:O80949; EMBL:AC004697; NID:G3402671; PID:G3402675
 A;Experimental source: cultivar Columbia
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A;Reference number: A84420; MUID:20083487; PMID:10617197
 A;Accession: H84815
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-189 <STO>
 A;Cross-references: GB:AE002093; NID:G3402675; PIDN:AAC28978.1; GSPDB:GN00139
 C;Genetics:
 A;Gene: T16B24.4; At2g39320
 A;Map position: 2
 A;Introns: 9/3; 34/3; 108/1
 Alignment Scores:
 Pred. No.: 0.573 Length: 189

Score: 99.00 Matches: 31
 Percent Similarity: 41.13% Conservative: 27
 Best Local Similarity: 21.99% Mismatches: 43
 Query Match: 6.03% Indels: 40
 DB: 2 Gaps: 5
 AF334735 (1-954) x T02570 (1-189)
 QY 220 AGAGCAACCGGACAAATATA-----CCAGCTTTTGCAGCAGCCTATTTTGAGAGC 270
 Db 80 GlnLysGluProaspLysValIleHisMetSerTyrLeuAlaGlyIleHisPheAsnSer 99
 QY 271 CTTCTAGAGAAA-----ACAGAGAAAACCAACTTTGATCCAGCAGAA 312
 Db 100 IleTyrLysLysAsnLysGluLysGlySerArgSerSerSerSerSerSerAlaVal 119
 QY 313 TGGGGAGTAGGTAGAGACCGCTTCTATACATCATGCTTCGAGAGAGCAGACCA 372
 Db 120 Trp---MetLysLeuGlnArgLysLysGluAsnGluAlaLysLysLysGluGluGlu 138
 QY 373 CCTGAGAAAAGTGATCCTAAACAAGAGAGTCTCAGATATCTGGGAAGGAGGAGAGACA 432
 Db 139 LysGluArgLysAspMetGluLysGluLysLys----- 150
 QY 433 TCAGTCACCATCTTAGACTCTTCTGAGGAAGATAAGAAAAGAGAGGTTGCTGCTGC 492
 Db 151 -----LysAspLysGluaspLysLysLysAsp----- 159
 QY 493 AAAATCCAGCTGCTTCGCGGACACATAGCCAGAGAGAGGCAAGAAATGAAAACA 552
 Db 160 -----LysGluaspLysLysAlaLysVal 168
 QY 553 AATAGTCTTCAAAATGAGGAAAAGAGGAAAACAAGTGGAGGACACTGTTTACCTCCAG 612
 Db 169 GlnLysGluLysLysGluLysLysLysLysAsnArgAsnHisPheHisTyrSer 188
 QY 613 GAA 615
 Db 189 Glu 189
 RESULT 34
 S44893
 ZK1236.7 protein - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 14-Sep-1994 #sequence_revision 12-May-1995 #text_change 23-Mar-2001
 C;Accession: S44893
 R;Favell, A.D.
 submitted to the EMBL Data Library, May 1993
 A;Description: Sequence of the C. elegans cosmid ZK1236.
 A;Reference number: S44622
 A;Accession: S44893
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-241 <FAV>
 A;Cross-references: EMBL:L13200; NID:G289748; PID:G289749
 C;Genetics:
 A;Introns: 122/3; 209/1
 Alignment Scores:
 Pred. No.: 0.591 Length: 241
 Score: 99.00 Matches: 49
 Percent Similarity: 40.93% Conservative: 39
 Best Local Similarity: 22.79% Mismatches: 94
 Query Match: 6.03% Indels: 33
 DB: 2 Gaps: 7
 AF334735 (1-954) x S44893 (1-241)
 QY 75 CGGAACCGCGCACCAGCTCGGAGAGAAAGAGGTTCCATAGGCAGTCTTACCAAGAA 134
 Db 3 ArgAsnAlaArgArgValAsnArg-----AspGlu 13
 QY 135 GATGTCGATTCTCCCAACACCACTACCGAATTCACCAAGGATTTGGGAATCTTCT 194

```

Db      14  GInGluAspGlyPheValAsnHisMetMetAsnAspGlyGluAspValGluAspLeuAsp 33
QY      195  TGAAGGCTGACACCGAGATTCTTGAGAGAGCAACCGGACATATATACCGACTTTTGCAGC 254
Db      34  GlyGlyAlaGluGlnPheGluTyrAsp-GluAspGlyLysIleGlyLysArgLysAl 53
QY      255  ACCCTATTTTGTAGAGCTTCTAGAGAAAGAGAGAAACCAACTTTGTATCCAGCAATG 314
Db      53  aAlaLysLeuGlnAlaLysGluGluLysArgGlnMetArgGluTyrGluValArgGluAr 73
QY      315  GGGAGTAGTAGAGACCGCTTCTATACAATCATGCTATCGAGGAGCAGCAACACCACC 374
Db      73  gGluGluArgLysArgArgGluGluGluArgGluLysLysArgAspGluGluArgAlaLys 93
QY      375  TGAGAAAAGTGTATCTTAAACAAGAGAGTCTCAGATATCTGGGAAGGAGGAGAGACATC 434
Db      93  sGluGluAlaAspGluLysAlaGluGluGluArgLeuArgLysGluArgGluGluGluGlu 111
QY      435  AGTCACCATCTTAGACTCTTCTGAGAGAGATAGCAAGAAAAGAGAGGTTGCTGTCTCAA 494
Db      112  -----LysGluArgLysGluHisGluGluTyrLeuAlaMetLys 124
QY      495  AATCAAGCTGCTTCCGGGGACATAGCCAGAGAGGAGGCAAGAAAATGAAAACAAA 554
Db      124  sAlaSerPheAlaLys-----GluGluGluGly-----ThrAs 135
QY      555  TAGTCTTCAAAATGAGGAAAAGAGCAAAACAAGTGGAGACACTGTTTTACCTCCAGGA 614
Db      135  pAlaIleGluGlyGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 154
QY      615  ACATGAAAATAATCCAAATCCATCAACCTCTTATTATGTCATTCTTCTCTGAGGAA 674
Db      154  sThrAsnLysValAlaAsnIleAspGluLeuSer-----SerHisPheGlyLeuLysE 172
QY      675  GCAAGATTTGTTGTGTAATAACATTCGTTACTGTTGTGA 717
Db      172  rGluAspAlaVal-----AsnArgLeuGlnHisPheIleGlu 184

RESULT 35
S55395
neurofilament protein M - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C:Accession: S55395
R:Vitadello, M.; Vettore, S.; Lamar, E.; Chien, K.R.; Gorza, L.
A:Description: Neurofilament mRNA and protein are expressed in precursors of heart condu
A:Reference number: S55395
A:Accession: S55395
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-644 <VT>
A:Cross-references: UNIPROT:P54938; EMBL:247378; NID:g854352; PIDN:CAA87454.1; PID:g8543
C:Superfamily: cytoskeletal keratin

Alignment Scores:
Pred. No.: 0.668 Length: 644
Score: 99.00 Matches: 48
Percent Similarity: 40.96% Conservative: 29
Best Local Similarity: 25.53% Mismatches: 70
Query Match: 6.03% Indels: 41
DB: 2 Gaps: 7

AF334735 (1-954) x S55395 (1-644)
QY      121  GTTCTTACCAAGAGATGTCGATTCCATCTCTCAACACCCTACCGAATCCACAGGA 180
Db      172  LeuLeuAsnValLysMetAlaLeuAspLysGluIleAlaLysArg----- 187
QY      181  TTTGGGAATCTCTTGAAGGG-----CTGACA 207
Db      188  -----LysLeuLeuGluGlyGluThrArgPheSerThrPheSerGlySerIleThr 205

```

```

QY      208  CGCAGATTCTGAGAGACCAACCGACATATATACCGCTTTTGACAGCAGCTATTTTGAG 267
Db      206  GlyProLeuTyrThrHisArgGlnProSerVal----- 216
QY      268  AGCTTCTAGAGAAAGAGAGAAAACCAACTTTGATCCAGCAGCAATGGGGAGTAAAGTA 327
Db      217  ThrIleSerSerLysIleGlnLysThrLysValGluAlaProLys-----LeuLysVal 234
QY      328  GAAGACCGCTTCTATAACAATCATGTCATTCGAGGAGCAAGAACCCCTCGAGAAAGTGT 387
Db      235  GlnHisLysPheValGluGluIleLeuGluThrLysValGluAspGluLysSerGlu 254
QY      388  CTTAAACA-----GAACAGTCTCAGATATCTGGGAAGGAGCAAGAG 429
Db      255  MetGluAspAlaLeuThrAlaIleAlaGluGluLeuAlaValSerValLysGluGlu 274
QY      430  ACATCAGTCACCATCTTAGACTCTTCTGAGGAAGATAGGAAAAGAGAGAGTGTCTGCT 489
Db      275  LysGluGluGluAlaGluGlyLysGluGluGluAlaGluGluGluValAlaAla 294
QY      490  GTCAAAATCCAA-----GCTGCTTCGGGGACACATAGCCAGAGAGAGGCAAGAAA 543
Db      295  AlaLysLysSerProValLysAlaThrThrProGluIleLysGluGluGluGlyGluLys 314
QY      544  ATGAAAACAATAGTCTTCAAAATGAGGAAAAGAGGAAAACAAGTGGAGGACACTGTTT 603
Db      315  GluGlu-----GluGlyGlnGluGluGluGluGluGluGluGluGluGlyVa 330
QY      604  TACCTCCAGGAAACATGAAAAA 625
Db      330  LysSerAspGlnAlaGluGlu 337

RESULT 36
A71623
probable secreted protein PF0115w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004
C:Accession: A71623
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743; PMID:9804551
A:Accession: A71623
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1192 <GAR>
A:Cross-references: UNIPROT:O96127; GB:AE001373; GB:AE001362; NID:g3845097; PIDN:AAC7181
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PF0115w

Alignment Scores:
Pred. No.: 0.721 Length: 1192
Score: 99.00 Matches: 38
Percent Similarity: 48.82% Conservative: 24
Best Local Similarity: 29.92% Mismatches: 42
Query Match: 6.03% Indels: 23
DB: 2 Gaps: 5

AF334735 (1-954) x A71623 (1-1192)
QY      277  GAGAAAAGAGAGAAAACCAACTTTGATCCAGCAAGATGGGGAGTAAAGTAGAAGACCGC 336
Db      416  GluAsnLysGluGluThrGluValAspGluLysLys-----ThrGluLysAlaGluGlu--- 433
QY      337  TTCTATAACATCATGTCATTCGAGAGCAACCACTCGAGAAAGTGTATCTTAAACAA 396
Db      434  -----GluLeuGluGluAspLysGluGluSerGluLysAspLys 446
QY      397  GAAGAGTCTCAGATATCTGGGAAGGAGGAGACATCAGTCACCATCTTAGACTCTTCT 456

```

Db 447 GluGluSerGlu-----LysAspLysGluGluSerGluLysAspLysGluGluSer 463
QY 457 GAGAAAGATAAGAAAGAAAGAGAGTTGCTGCTCAAAATCCAAAGCTGCTCCGGGA 516
Db 464 GluLysAspLysGluLysThrGluGluAspGluGluLysThrGluLysGlyThr 483
QY 517 CACATAGCCAGAGAGAGGCA-----AAGAAATGAA----- 549
Db 484 GluValTyrLysLysGluThrAspValAspGluLysLysGluLysGlyGluGlu 503
QY 550 ----ACAATAGTCTTCAAAATGAGAAAGAGAAACACAGTGCAGACACTGTTTAC 606
Db 504 GlyThrAspGluGluAspLysGluLysGluGluAsp-AspGluGluThrLysValG 523
QY 607 CTCAGGAAACATGAAAA 625
Db 523 uGluLysLysThrGluLys 529

RESULT 37
S45706
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 10-Dec-1994 #sequence_revision 19-Apr-1996 #text_change 24-Nov-1999
C:Accession: S45706
R: Peng, M.; Fan, H.; Kirtley, T.L.; Caswell, A.H.; Schwartz, A.
FEBS Lett. 348, 17-20, 1994
A: Title: Structural diversity of triadin in skeletal muscle and evidence of its existenc
A: Reference number: S45704; MUID: 94298946; PMID: 8026576
A: Accession: S45706
A: Status: preliminary
A: Molecule type: mRNA
A: Residues: 1-213 <PEN>
C: Superfamily: histone H1

Alignment Scores:
Pred. No.: 0.648 Length: 213
Score: 98.50 Matches: 48
Percent Similarity: 44.39% Conservative: 35
Best Local Similarity: 25.67% Mismatches: 71
Query Match: 6.00% Indels: 33
DB: 2 Gaps: 7

AF334735 (1-954) x S45706 (1-213)
QY 41 GAAAAACGGTTACCCAGCAACTAGAAAAACACCGAACCGCGGCACAGCTCGGAGA 100
Db 19 GluLysGlnAlaLysAlaLysIleGluLysGluLysGluLysValSerAla--AlaSerThrL 38
QY 101 GAAAGAGGTTCCATAGGCAGTTCTTACCAGAGAGATGTCGATTCCATTCTCCACACCC 160
Db 38 yLysAlaValProAlaLysLysGluLysThrLysThrValGluGluThrA 58
QY 161 ACTACGAATTCACAGAGATTGGGAATCTTTGAAGGGCTGACACCGAGATTCTGA 220
Db 58 rGlyGluLysProGlyLysIleSerSerValLysLys-----AspLysGluLysThrL 76
QY 221 GAGAGCAACCGGACAAATATACAGCTTTTCAGCAGCCTTTTTCAGCAGCCTTTTTCAGAGCCTTCTAGAGA 280
Db 76 yGluLysGluValLysValProAla-----SerLeuLysGluL 89
QY 281 AAGAGAGAAAAACCACTTTGATCCAGCAGATGGGGAGTAGAGTGAAGACCGCTTCT 340
Db 89 yGlySerGluThrLysLysAsp-----GluLysThrSerLysProGluProGlnIleL 107
QY 341 ATACAAATCATGCTTCGAGGAGCAAGACCACTGAGAAAGTGTATCTTAAACAGAG 400
Db 107 yS-----LysGluLysProGlyLysGluLysValLysProLysProProG 122
QY 401 AGTCTCAGATATCTGGAGAGGAGAA-----GAGACATCAGTCACCA 442
Db 122 InProGlnIleLysLysGluGluLysProGluGlnAspIleMetLysProGluLysThrA 142
QY 443 TCTTAGACTCTTCTGAGGAAGATAAGAAAAAGAGAGGTTGCTGCTGCTCAAAATCCAAG 502

Db 142 laLeuHisGlyLysProGluGluLysValLeuLysGlnValLysAlaValThrThrGluL 162
QY 503 CTGCTTCCGGGACACATAGCCAGAGAGGAGCAAGAAATGAAACAAATAGTCTTC 562
Db 162 yS-----HisValLysProLysProAlaLysLys-----AlaG 173
QY 563 AAAATGAGAAAGAG 579
Db 173 luHisGlnGluLysGlu 178

RESULT 38
T15827
hypothetical protein C53C9.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000
C:Accession: T15827
R: Bentley, D.
submitted to the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans cosmid C53C9.
A:Reference number: Z18413
A:Accession: T15827
A:Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-418 <BEN>
A: Cross-references: EMBL: U28734; NID: g861255; PID: g861256; PIDN: AAB52603.1; GSPDB: GN000
A: Experimental source: strain Bristol N2; clone C53C9
C: Geneticks:
A: Gene: CESP: C53C9.2
A: Map position: X
A: Introns: 25/3; 187/1; 310/2; 380/3; 405/3

Alignment Scores:
Pred. No.: 0.704 Length: 418
Score: 98.50 Matches: 45
Percent Similarity: 40.31% Conservative: 32
Best Local Similarity: 23.56% Mismatches: 73
Query Match: 6.00% Indels: 41
DB: 2 Gaps: 7

AF334735 (1-954) x T15827 (1-418)
QY 76 GGAACCGCGGCACCAGCTCGGAGAGAAAGAGGTTCCATAGGCAGTTCTTACCAGAG 135
Db 244 GlySerGlyGlyPheLeuLysValArgAspValLeuProHisThrValGlyGlyLysAsp 263
QY 136 ATGTCG-----ATTCATTC-----TCCACACCCAC 162
Db 264 ileGluGluGluLysSerGluGlyIleValProLeuGlnSerGlyThrAsn 283
QY 163 TACCGAATTCACAA-----GGATTGGG-----AATCTTCTT 195
Db 284 LysLeuAlaSerGlnArgGlyMetThrGlyPheGlyThrProArgAsnThrGlnLeuArg 303
QY 196 GAAAGGCTGACACGCGAGATTCTGAGAGCAACCGGACAAATATACCAGCTTTTCGAGCA 255
Db 304 AlaGlyTyrLysGluTrpIleGluAsp----- 313
QY 256 GCCTATTTTGAAGCGCTTCTAGAGAAAGAGAGAAACCACTTGTATCCAGCAGANTGG 315
Db 314 -----TyrGluAlaLeuLysGluTrpGluGluThrLys-----ProGlySer 329
QY 316 GGGAGTAGGTAGAGACCGCTTCTATAACATCATGCTTCGAGAGCAAGAACACACCT 375
Db 330 AlaSerSerValAspProPheGlyHisTyrLysLysLysPheGluGluArgGlu----- 347
QY 376 GAGAAAGTGCCTCTAAACAGAGAGTCTCAGATATCTGGGAGAGAGAGAGAGCATCA 435
Db 348 -----SerSerArgGlnSerGluIleAspSerGlnSerValLysAlaSer 362
QY 436 GTCACATCTTAGACTCTTCTGAGAGAGATAAGAAAAAGAGAGGTTCTGCTGTCATAA 495
Db 363 GluProValGluProGluGluGluGluGluGluGluGluGluGluGluGluGlu 382


```

QY 368 AACACCTGAGAAAAGTGTCTTAAACAAGAGAGTCTCAG-----A 409
Db 760 luthrlysgluGluGlyLysProLeuGlnGlnGluLysGluLysGluLysGluGlyG 780
QY 410 TATCTGGGAGGAGGAGAGACATCAGTACCATCTTACACTCTTCTGAGGAGATAAG 469
Db 780 luGlyGlySerGluGluGluGlySer-----AspLysGlyAlaLysGlySerA 796
QY 470 AAAAAGAGAGGTTGCTGCT-----GTCAAAATCCAAG 502
Db 796 rglYsgluAspIleAlaValAsnGlyGluValGluGlyLysGluGluValGluGlnGluT 816
QY 503 CTCCTCTCCGGGACACATAGCCAGAGAGAGCAAGAAATGAAACAAATAGTCTTC 562
Db 816 hrlYsgluLysGlySer---GlyArgGluGluGluGlyValValThrAsnGlyLeuA 835
QY 563 AAAATGAGGAAAAGAGGAAAACAACAGTGAGGA 594
Db 835 splEusSerProAlaAspGluLysGlyGly 845

RESULT 41
S60254
nuclear receptor co-repressor N-CoR - mouse
C:Species: Mus musculus (house mouse)
C>Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C:Accession: S60254
R:Hoerlein, A.J.; Naeae, A.M.; Heinzel, T.; Torchia, J.; Gloss, B.; Kurokawa, R.; Ryan,
Nature 377, 397-404, 1995
A:Title: Ligand-independent repression by the thyroid hormone receptor mediated by a nu
A:Reference number: S60254; MUID:96008539; PMID:7566114
A:Accession: S60254
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-2453 <HOB>
A:Cross-references: UNIPROT:Q60974; EMBL:U35312; NID:g1022717; PIDN:AAB17125.1; PID:g102

Alignment Scores:
Pred. No.: 0.977 Length: 2453
Score: 98.00 Matches: 45
Percent Similarity: 39.34% Conservative: 27
Best Local Similarity: 24.59% Mismatches: 63
Query Match: 5.97% Indels: 48
DB: 2 Gaps: 9

AF334735 (1-954) x S60254 (1-2453)
QY 67 AAAACAACCGAACCGCGCCAGCTCGAGAGAAAGAGGTTCATAGGAGT--- 123
Db 457 LysAsnPhenGlyLeuIleAlaSerTyrLeuGluArgLysSerValProAspCysValLeu 476
QY 124 -----CTTACCAGAGATGTCGATTCCATTCTCAACACCCACTACCGA----- 168
Db 477 TyrTyrLeuThrLysLys-----AsnGluAsnTyrLysAlaLeu 490
QY 169 ATTCCAAGAGATTTCGGAATCTTCTTGAAGGCTCACACGGAGATTCTGAGAGACAA 228
Db 491 ValArgAsnTyrGlyLys---ArgArgGlyArgAsnGlnIleAlaArgProSer 509
QY 229 CCGACAAATATACAGCTTTTGCGAGCGCTATTATTGAGAGCCCTCTTAGAGAAAGAGAG 288
Db 510 GlnGlu-----GluLysValGlu 515
QY 289 AAAACCAACTTTCATCCAGCAGAAATGGGAGTAAGTAGAGACCGCTTCTATAACAAT 348
Db 516 GluLysGluGluAspLysAlaGluLysThrGluLysGlu----- 529
QY 349 CATGCATTTCGAGGAGCAACACCTGAGAAAGTATCCTTAACAAGAGAGTCTCAG 408
Db 530 -----GluGluLysLysAspAspGluGluLysAspAspLysGluAspSerLysGlu 546
QY 409 ATATCTGGGAGGAGAGAGACATCAGTCACCATCTTACACTCTTCTGAGGAGATAG 468
Db 547 ThrThrLysLysAspArgThrGluAlaThr-----AlaGluGluProGlu 562

```

```

QY 469 GAAAAGAGAGAGTGTCT-----GCTGTCAAAATCCAAGCTCGCTTCCGG 513
Db 563 GluArgGluGlnValThrProArgGlyArgLysThrAlaAsnSerGlnGlyArgGlyLys 582
QY 514 GGACACATAGCCAGA-----GAGGAGGCAAGAAATGAAACAAATAGTCTTCAA 564
Db 583 GlyArgValThrArgSerMetThrSerGluAlaAlaAlaAlaAsnAlaAlaAlaAla 602
QY 565 AATGAGGAA 573
Db 603 ThrGluGlu 605

RESULT 42
S49313
protein kinase - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C>Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 16-Aug-2004
C:Accession: S52076; S49313
R:Wetterauer, B.W.; Hamker, U.; von Haeseler, A.; MacWilliams, H.K.; Simon, M.N.; Veron,
Biochim. Biophys. Acta 1265, 97-101, 1995
A:Title: A protein kinase from Dictyostelium discoideum with an unusual acidic repeat c
A:Reference number: S52076; MUID:95161460; PMID:7857991
A:Accession: S52076
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1094 <WR2>
A:Cross-references: UNIPROT:Q23915; EMBL:Z37981; NID:g551445; PIDN:CAA86053.1; PID:g551
C:Genetics:
A:Introns: 35/3; 104/1; 166/2
C:Superfamily: protein kinase homology
C:Keywords: ATP
F:20-281/Domain: protein kinase homology <KIN>
F:28-36/Region: protein kinase ATP-binding motif

Alignment Scores:
Pred. No.: 0.984 Length: 1094
Score: 97.50 Matches: 44
Percent Similarity: 41.38% Conservative: 40
Best Local Similarity: 21.67% Mismatches: 64
Query Match: 5.94% Indels: 55
DB: 2 Gaps: 8

AF334735 (1-954) x S49313 (1-1094)
QY 38 AAGAAAACCGTTTACCCAGCACTAGAAAACACCGGACCGCGCGCCAGCTCG 97
Db 558 GluGluGluGluTyrGlnArgGlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGln 577
QY 98 AGAGAAAGGAGGTTCCATAGGCGAGTCTTAC----- 128
Db 578 GlnGlnGlnHisGlnHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 597
QY 129 -CAAGAGAGATGTCGATTCTTCATCTTCAACACCCTACCGAATTTCCACAGGATTGGGA 187
Db 597 uGluGluGluTyrAspAspValIleArgHis----- 607
QY 188 ATCTTCTTGAAGGCTGACCGCGAGATTCTCAGAGAGCAACCGGACCAATATACAGCTT 247
Db 608 -----AspThrAspSerGluGlu-GluSerLysAspLysThrProLeuP 622
QY 248 TTGCAGCAGCCTTATTTTGAGAGCCTTTAGAGAAAAGAGAGAAAACCAACTTTGATCCAG 307
Db 622 roTTPAspGlnHisPheGlu-----LysGlnLysGluSerGluAsn----- 635
QY 308 CAGAATGGGGAGTAGTAGTAGAAGACCGCTTCTATAAC-----ATCATGCAT 355
Db 636 -----LysValGluGlnGluGluThrAsnValValValAlaAsnSerGlnG 651
QY 356 TCGAGGAGCAAGAACCCCTGAGAAAAGTGTCTTAAACAAGAGAGTCTCAGATATCTG 415
Db 651 luThrAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 671

```


R:Johnson, D.; Wamsley, P.; Bradshaw, H.
submitted to the EMBL Data Library, February 1997
A:Description: The sequence of C. elegans cosmid ZK354.
A:Reference number: Z20120
A:Accession: T25991
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-224 <JOH>
A:Cross-references: UNIPROT:P91567; EMBL:U88172; PIDN:AB42259.1; GSPDB:GN00022; CESP:ZK
A:Experimental source: strain Bristol NZ; clone ZK354
C:Genetics:
A:Gene: CESP:ZK354.7
A:Map position: 4
A:Introns: 25/3

Alignment Scores:
Pred. No.: 1 Length: 224
Score: 96.50 Matches: 35
Percent Similarity: 45.24% Conservativeness: 22
Best Local Similarity: 27.78% Mismatches: 56
Query Match: 5.88% Indels: 13
DB: 2 Gaps: 4

AF334735 (1-954) x T25991 (1-224)

QY 229 CCGGCAATATACAGCTTTTCAGCAGCGCTATTTTGAGCGCTTCTAGAGAAAGAGAG 288
DB 92 ProGluAsnAsnLysGlnTyRPhelaPheTyTyR-----Ile 104
QY 289 AAAACCAACTTTGAT---CCAGCAGAA-----TGGGGAGTAAAGGTAGAGAC----- 333
DB 105 LysThrLysGluAspGlyProAlaArgAspAlaTrpLysAsnHisLysProAspGlyMet 124
QY 334 ---CGCTTCTATAACAATCATCATTCAGGAGCAAGAACCACTGAGAAAGTATCCT 390
DB 125 LysArgIleTyRLeSerPheLysGlyAspGluLysGlyAspGluLysLysAspGlu 144
QY 391 AAACAAGAGAGCTCTCAGATATCTGGGAGGAGGAGAGACATCAGTCACCATCTTAGAC 450
DB 145 LysLysGluLysLysGluLysGluLysGluLysGluLysGluLysGluLysLysAsp 164
QY 451 TCTTCTGAGGAGATAGAGAAAGAGAGGTTCGTCTGTCTCAAAATCCAAAGCTCCCTTC 510
DB 165 LysLysGluAlaLysLysGluLysLysGluLysGluLysGluLysLysGluLys 184
QY 511 CCGGACACATGACGAGAGAGCAAGAAAGATGAAACAAATAGTCTTCAAAATGAG 570
DB 185 LysLysGluLysGluLysGluLysGluLysGluLysGluLysGluLysGluLys 204
QY 571 GAAAGAGAGGAAACAAAG 588
DB 205 LysLysGluLysLys 210

RESULT 48
A45605
mature-parasite-infected erythrocyte surface antigen MESA - malaria parasite (Plasmodium
C:Species: Plasmodium falciparum
C:Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A45605; A54517
R:Coppel, R.L.
Mol. Biochem. Parasitol. 50, 335-347, 1992
A:Title: Repeat structures in a Plasmodium falciparum protein (MESA) that binds human ex
A:Reference number: A45605; MUID:92158014; PMID:1741020
A:Accession: A45605
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1526 <COP>
A:Cross-references: UNIPROT:P08116
A:Experimental source: Papua New Guinean isolate FC27
A>Note: sequence extracted from NCBI backbone (NCBI:83648, NCBI:83656)
R:Coppel, R.L.; Culvenor, J.G.; Bianco, A.E.; Cretzner, P.E.; Stahl, H.D.; Brown, G.V.;
Mol. Biochem. Parasitol. 20, 265-277, 1986
A:Title: Variable antigen associated with the surface of erythrocytes infected with matu

A:Reference number: A54517; MUID:87014571; PMID:3531849
A:Accession: A54517
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 222-443, 'K', 445 <CO2>
A:Cross-references: GB:M15319; NID:g160060; PID:g552170
C:Superfamily: ring-infected erythrocyte surface antigen; dnaJ amino-terminal homology
C:Keywords: surface antigen; tandem repeat

Alignment Scores:
Pred. No.: 1.27 Length: 1526
Score: 96.50 Matches: 46
Percent Similarity: 41.15% Conservativeness: 33
Best Local Similarity: 23.96% Mismatches: 57
Query Match: 5.88% Indels: 56
DB: 2 Gaps: 8

AF334735 (1-954) x A45605 (1-1526)

QY 42 AAAAAGCGTTTACCCAGCACTAGAAAAACAACCGGCGGCGCACCAGCTCGGAGAG 101
DB 874 LysLysArgValLysLys-----ArgAsnAsnLysAsnGluArg----- 886
QY 102 AAGGAGGTTTCCATAGGCGAGTTCT---TACCAAGAAGATGTCGATTTCATTCTCCAAAC 158
DB 887 LysAspAsnValLysGluLysGluLysGluLysGluLysGluLysGluLysGluLys 902
QY 159 CCATACCGAATCCACAGGATTTGGGAATCTTCTTGAAGGGCTGACACCGAGATTCT 218
DB 903 -----LysAspThrAlaAsnLysAsp-LysGluLysGluLysGluLys 913
QY 219 GAGAGCAACCGGCAATATATACAGCTTTTTCAGCAGCGCTATTTTGAGAGCGCTTCTAGA 278
DB 913 uGlnGluLysGluLys-----GluGluValLysGluLysGluLysGluLysGluLysGluLys 923
QY 279 GAAAGAGAGAAAAACCAACTTTTGATCCAGCAGAAATGGGGAGTAGAGTAGAGACCGCTT 338
DB 923 uLysGluValLysGluLysGluLysGluLysGluLysGluLysGluLysGluLysGluLys 942
QY 339 CTATAACATCATCGATTCGAGGAGCAAGAACCACTGAGAAAAGTAGATCTCTAAACAAGA 398
DB 943 -----GluGluValLysGluLysGluLysGluLysGluLysGluLysGluLysGluLys 960
QY 399 AGAGTCTCAGATATCTGGGAGGAGGAGAGACATCAGTCACCATCTTAGACTCTTCTGA 458
DB 960 AspThrGluSerLysAspLysGlu-----IleGlu 970
QY 459 GGAAGATAAGAAAAAGAGAGGTTGCTGCTGTCTCAAAATC----- 498
DB 970 uGlnGluLysGluLysGluLysGluLysGluLysGluLysGluLysGluLysGluLysAsp 990
QY 499 -----CAAGCTCGCTTCGGGGGACATAGCCAGAGAGGAGGAGGAGGAGGAGGAGGAG 545
DB 990 pLysValLysGluLysGluLysGluLysGluLysGluLysGluLysGluLysGluLysGluLys 1010
QY 546 GAAACAAATAGTCTTCAAAATAGGAAAAAGAG 579
DB 1010 lLysLysArgAsnAsnLysAsnGluLysGluLysGluLysGluLysGluLysGluLys 1021

RESULT 49

B64171
ABC-type transport protein uup-2 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 16-Aug-2004
C:Accession: B64171
R:Flaischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, J.
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
Science 269, 496-512, 1995
A:Authors: Ghem, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: B64171

	:::		:::	
Db	511	AlaLysGluGluGluGlyGlyGluInGlyGluGlnGlyGluThrLysGluAlaGlu	530	:::
Qy	568	GAGCAAAAGAGGAA	582	
Db	531	GluGluLysLysAsp	535	

Search completed: February 15, 2005, 10:07:41
Job time : 61.5 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 15, 2005, 09:49:49 ; Search time 161.5 Seconds
(without alignments)
6049.827 Million cell updates/sec

Title: AF334735
Perfect score: 1641
Sequence: 1 TCGCCCTTCTCGCGCGCG.....AAAAAAAAAAAAAAAAAAAA 954

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-O=/cn2.1/USPTO spool_p/AF334735/runat_15022005_094942_10092/app_query.fasta.1.1095
-DB=UniProt_03 -QFMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=AF334735 @CGN.1.1.244 @runat_15022005_094942_10092 -MCPUS=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DSV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 03:1
1: uniprot_sprot:1
2: uniprot_trembl:1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	812	49.5	163	1 SP17 PAPA	Q95230 papio hamad
2	778	47.4	151	1 SP17 HUMAN	Q15506 homo sapien
3	758	46.2	151	1 SP17 MACPA	Q19021 macaca fasc
4	710	43.3	143	2 Q9BDQ6	Q9bdq6 macaca mula
5	684.5	41.7	212	2 Q9XT29	Q9xt29 callithrix
6	589	35.9	147	2 Q9TU08	Q9tu08 ovula aries
7	569	34.7	146	1 SP17 RABIT	P36425 oryctolagus
8	561.5	34.2	148	2 Q9ZIK2	Q9zik2 rattus norv
9	558	34.0	149	1 SP17 MOUSE	Q62252 mus musculu
10	544	33.2	153	1 SP17 WACEU	Q62770 macropus eu
11	502.5	30.6	179	1 SP17 MONDO	Q62771 monodelphis
12	491.5	30.0	141	2 Q6SJ92	Q6sj92 ornithorhyn
13	364	22.2	495	2 Q6DIJ9	Q6dij9 xenopus tro
14	134	8.2	189	2 Q20199	Q20199 caenorhabdi
15	131	8.0	422	2 Q6WN93	Q6wn93 brachydanio
16	122.5	7.5	368	2 Q25114	Q25114 hemiceurot

17	122.5	7.5	369	1	KAPR_STRPU	Q26619 strongyloce
18	121.5	7.4	379	2	Q9HAY3	Q9hay3 homo sapien
19	121.5	7.4	379	2	Q9HAY5	Q9hay5 homo sapien
20	121.5	7.4	493	2	Q75952	Q75952 homo sapien
21	121	7.4	596	2	Q9GM47	Q9gm47 macaca faec
22	119.5	7.3	562	2	Q8BU41	Q8bu41 debaryomyce
23	118.5	7.2	382	2	Q9BUB1	Q9bub1 homo sapien
24	118.5	7.2	403	1	KAP2 HUMAN	P13861 homo sapien
25	116.5	7.1	216	2	Q7S160	Q7s160 neurospora
26	116.5	7.1	401	2	Q8K1M2	Q8k1m2 rattus norv
27	116.5	7.1	498	2	Q710D7	Q710d7 vulpes vulp
28	115.5	7.0	400	1	KAP2 BOVIN	P00515 bos taurus
29	115	7.0	532	2	Q9T069	Q9t069 arabidopsis
30	114	6.9	992	2	Q7S5R9	Q7s5r9 neurospora
31	113.5	6.9	443	2	Q9LYR0	Q9lyr0 arabidopsis
32	112.5	6.9	3901	2	Q8N533	Q8n533 caenorhabdi
33	112	6.8	845	2	Q8NE71	Q8ne71 homo sapien
34	112	6.8	3290	2	Q7RRX0	Q7rrx0 plasmodium
35	111.5	6.8	396	2	Q6TU32	Q6tu32 aplysia cal
36	111.5	6.8	446	2	Q9LVA0	Q9lva0 arabidopsis
37	111	6.8	402	2	Q8K1M3	Q8k1m3 mus musculu
38	110.5	6.7	134	2	Q9CYV3	Q9cyv3 m mus muscu
39	110.5	6.7	921	2	Q8QTD2	Q8qtd2 white spot
40	110.5	6.7	922	2	Q8VAW5	Q8vaw5 white spot
41	110.5	6.7	922	2	Q9ILB5	Q9ilb5 white spot
42	110	6.7	103	2	Q9SD11	Q9sd11 arabidopsis
43	109	6.6	312	2	Q6CCN1	Q6ccn1 yarrowia li
44	108	6.6	403	1	KAPR_BLAEM	P31320 blastoclad
45	108	6.6	1132	1	YKK5_YEAST	P34250 saccharomyc

ALIGNMENTS

RESULT 1

SP17_PAPHA

ID SP17_PAPHA STANDARD; PRT; 163 AA.

AC Q95230;

DT 01-NOV-1997 (Rel. 35, Last created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Sperm surface protein SP17 (Sperm autoantigenic protein 17).

GN Name=SP17; Synonyms=SP17;

OS Papio hamadryas (Hamadryas baboon).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;

OC Cercopitheciinae; Papio.

OX NCBI_TaxID=9557;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Testis;

RX MEDLINE=97264490; PubMed=9110316;

RX DOI=10.1002/(SICI)1098-2795(199705)47:1-66::AID-MRD9>3.0.CO;2-O;

RA Adoyo P.A., Lea I.A., Richardson R.T., Widgeon E.E., O'Rand M.G.;

RT "Sequence and characterization of the sperm protein SP17 from the

baboon."

RL Mol. Reprod. Dev. 47:66-71(1997).

CC -!- FUNCTION: Sperm surface zona pellucida binding protein. Helps to

bind spermatozoa to the zona pellucida with high affinity. Might

function in binding zona pellucida and carbohydrates (By

similarity).

CC -!- SUBUNIT: Homodimer (By similarity).

CC -!- SUBCELLULAR LOCATION: Membrane-associated (Potential).

CC -!- TISSUE SPECIFICITY: Testis- and sperm-specific.

CC -!- SIMILARITY: Contains 1 IQ domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@sib-sib.ch).

```
DR EMBL; U75209; AAB38534.1; -.
DR HSSP; P12367; 1R2A.
DR InterPro; IPR000048; IQ region.
DR InterPro; IPR003117; RIIA.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF02197; RIIA; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00394; RIIA; 1.
DR PROSITE; PS50096; IQ; 1.
KW Membrane.
FT DOMAIN 114 143 IQ.
SQ SEQUENCE 163 AA; 18755 MW; A890B59B3F6469BD CRC64;

Alignment Scores:
Pred. No.: 4 35e-61 Length: 163
Score: 812.00 Matches: 158
Percent Similarity: 98.1% Conservative: 3
Best Local Similarity: 96.34% Mismatches: 2
Query Match: 49.48% Indels: 1
DB: 1 Gaps: 0

AF334735 (1-954) x SP17_PAPHA (1-163)
QY 136 ARTGCGATTCCATTCTCCACACCCACTACCGAATTCACAGGATTTGGGAATCTTCTT 195
DB 1 MetSerIleProPheSerAsnThrHisTyArgIleProGlnGlyPheGlyAsnLeuLeu 20
QY 196 GAAGGCTGACAGCGGAGATTCTGAGAGACACCGGACATATACACGCTTTTCAGCA 255
DB 21 GluGlyLeuThrArgGluLeuArgGluGlnProAspAsnIleProAlaPheAlaAla 40
QY 256 GCCTATTGAGAGCTTCTAGAGAAAGAGAGAAACCACTTTGATCCAGCAGAAATGG 315
DB 41 AlatyPheGluSerLeuLeuGluysArgGluysThrAsnPhAspProAlaGluTrp 60
QY 316 GGGAGTAAAGTGAAGACCGCTTCTATACCAATCATGCTTCGAGAGAGCAAGAACCT 375
DB 61 GlySerLysValGluAspArgPheTyAsnAsnHisAlaPheGluGluGlnGluProPro 80
QY 376 GAGAAAGTATCTTAACAAGAGCTTCAGATATCTGGAGAGAGAGAGAGATCA 435
DB 81 GluLysSerAspProLysGlnGluSerGlnValSerGlyLysGluGluThrSer 100
QY 436 GTCACCATCTAGACTTCTCAGAGAGATAGGAAAGAGAGAGAGAGAGAGAGAGAG 495
DB 101 ValThrIleLeuAspSerSerGluGluAspLysGluLysGluGluValAlaAlaVallys 120
QY 496 ATCCAGCTGCCTTCGCGGGACACATAGCCAGAGAGAGAGAGAGAGAGAGAGAGAG 555
DB 121 IleGlnAlaAlaPheArgGlyHisValAlaArgGluGluValLysLysMetLysThrAsp 140
QY 556 AGCTTTCAAAATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 615
DB 141 SerLeuGlnAsnGluGluLysGluGluAsn-SerGluAspThrGlyPheThrSerArgTh 160
QY 616 ACATCAAAA 625
DB 160 rHisGluLys 163

RESULT 2
SP17_HUMAN
ID SP17_HUMAN STANDARD; PRT; 151 AA.
AC Q15506; Q9BXF7;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Sperm surface protein Sp17 (Sperm autoantigenic protein 17) (Sperm
DE protein 17) (Sp17-1).
GN Name=SP17; Synonyms=SP17;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
```

```

RN RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=96305346; PubMed=8688458; DOI=10.1016/0167-4781(96)00077-2;
RA Lea I.A., Richardson R.T., Widgren E.E., O'Rand M.G.;
RT "Cloning and sequencing of cDNAs encoding the human sperm protein,
RT Sp17.";
RL Biochim. Biophys. Acta 1307:263-266(1996).
RN RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22280349; PubMed=12393185; DOI=10.1016/S0167-4781(02)00478-5;
RA Buchli R., De Jong A., Robbins D.L.;
RT "Genomic organization of an intron-containing sperm protein 17 gene
RT (Sp17-1) and an intronless pseudogene (Sp17-2) in humans: a new
RL Biochim. Biophys. Acta 1578:29-42(2002).
RN RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC FUNCTION: Sperm surface zona pellucida binding protein. Helps to
CC bind spermatozoa to the zona pellucida with high affinity. Might
CC function in binding zona pellucida and carbohydrates (By
CC similarity).
CC -! SUBUNIT: Homodimer (By similarity).
CC -! SUBCELLULAR LOCATION: Membrane-associated (Potential).
CC -! TISSUE SPECIFICITY: Testis- and sperm-specific.
CC -! SIMILARITY: Contains 1 IQ domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z48570; CA88459.1; -.
CC EMBL; AF334735; AAK20878.1; -.
CC EMBL; AF334810; AAK28125.1; -.
CC EMBL; BC032457; AAK32457.1; -.
CC FIR; J38243; J38243.
CC HSSP; P12367; 1R2A.
CC Genew; HGNC:11210; SPAL7.
CC H-InvdB; HIX0010221; -.
CC MIM; 608621; -.
CC GO; GO:0007338; P:fertilization (sensu Animalia); TAS.
CC GO; GO:0007283; P:spermatogenesis; TAS.
CC InterPro; IPR000048; IQ region.
CC InterPro; IPR003117; RIIA.
CC Pfam; PF00612; IQ; 1.
CC Pfam; PF02197; RIIA; 1.
```

```

DR SMART; SM00015; IQ; 1.
DR SMART; SM00394; RIIA; 1.
DR PROSITE; PS50096; IQ; 1.
KW Membrane.
FT DOMAIN 114 143 IQ.
SQ SEQUENCE 151 AA; 17406 MW; BC99EC310FA0E54A CRC64;

Alignment Scores:
Pred. No.: 3,62e-58 Length: 151
Score: 778.00 Matches: 151
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 47.41% Indels: 0
DB: 1 Gaps: 0

AF334735 (1-954) x SP17_HUMAN (1-151)

QY 136 ATGTCGATTCATTCCTCAACACCCACTACCAAGATTCACAGATTTGGGATCTTCTT 195
Db 1 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeu 20

QY 196 GAAGGGCTGACCGCAGATTCTGAGAGCAACCGGACAAATATACCACTTTTTCAGCA 255
Db 21 GluGlyLeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAla 40

QY 256 GCCTATTTTGGAGCCCTTCTAGAGAAAGAGAAACCAACTTTGATCCAGCAATCG 315
Db 41 AlaTyrPheGluSerLeuGluLeuArgGluGlnProAspAsnIleProAlaPheAla 40

QY 316 GAGAGTAAAGTAGAGACCGCTTCTATACCAATCATGATTCGAGAGCAAGAACCCACT 375
Db 61 GlySerLeuValGluAspArgPheTyrAsnAsnHisAlaPheGluGluGlnPro 80

QY 376 GAGAAAGTGATCTTAAACAAGAGAGTCTCAGATATCTGGGAGGAGAGAGATCA 435
Db 81 GluLysSerAspProLysGlnGluSerGlnIleProGlyLysGluGluGluAlaSer 100

QY 436 CTCACATCTTAGACTCTTCTGAGAGAGTAAAGAGAGAGAGAGAGAGAGAGAGAG 495
Db 101 ValThrIleLeuAspSerSerGluGluAspLysGluLysGluValAlaAlaValLys 120

QY 496 ATCCAAAGCTGCTTCCGGGGACACATAGCCAGAGAGGAGGCAAGAAATGAAACAAT 555
Db 121 IleGlnAlaAlaPheArgGlyHisValAlaAlaArgGluGluValLysMetLysThrAsp 140

QY 556 AGCTCTTCAAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 588
Db 141 SerLeuGlnAsnGluGluLysGluGluAsnLys 151

RESULT 4
Q9BDQ6 PRELIMINARY; PRT; 143 AA.
ID Q9BDQ6
AC Q9BDQ6;

similarity).
-!- SUBUNIT: Homodimer (By similarity).
-!- SUBCELLULAR LOCATION: Membrane-associated (Potential).
-!- TISSUE SPECIFICITY: Testis- and sperm-specific.
-!- SIMILARITY: Contains 1 IQ domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AF005551; AAB62888.1; -.
DR HSP; P12367; IR2A.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR003117; RIIA.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF02197; RIIA; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00394; RIIA; 1.
DR PROSITE; PS50096; IQ; 1.
KW Membrane.
FT DOMAIN 114 143 IQ.
SQ SEQUENCE 151 AA; 17401 MW; E681586AB713F03C CRC64;

Alignment Scores:
Pred. No.: 1,89e-56 Length: 151
Score: 758.00 Matches: 146
Percent Similarity: 98.01% Conservative: 2
Best Local Similarity: 96.69% Mismatches: 3
Query Match: 46.19% Indels: 0
DB: 1 Gaps: 0

AF334735 (1-954) x SP17_MACFA (1-151)

QY 136 ATGTCGATTCATTCCTCAACACCCACTACCAAGATTCACAGATTTGGGATCTTCTT 195
Db 1 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeu 20

QY 196 GAAGGGCTGACCGCAGATTCTGAGAGCAACCGGACAAATATACCACTTTTTCAGCA 255
Db 21 GluGlyLeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAla 40

QY 256 GCCTATTTTGGAGCCCTTCTAGAGAAAGAGAAACCAACTTTGATCCAGCAATCG 315
Db 41 AlaTyrPheGluSerLeuGluLeuArgGluGlnProAspAsnIleProAlaPheAla 40

QY 316 GAGAGTAAAGTAGAGACCGCTTCTATACCAATCATGATTCGAGAGCAAGAACCCACT 375
Db 61 GlySerLeuValGluAspArgPheTyrAsnAsnHisAlaPheGluGluGlnPro 80

QY 376 GAGAAAGTGATCTTAAACAAGAGAGTCTCAGATATCTGGGAGGAGAGAGATCA 435
Db 81 GluLysSerAspProLysGlnGluSerGlnIleProGlyLysGluGluGluAlaSer 100

QY 436 CTCACATCTTAGACTCTTCTGAGAGAGTAAAGAGAGAGAGAGAGAGAGAGAGAG 495
Db 101 ValThrIleLeuAspSerSerGluGluAspLysGluLysGluValAlaAlaValLys 120

QY 496 ATCCAAAGCTGCTTCCGGGGACACATAGCCAGAGAGGAGGCAAGAAATGAAACAAT 555
Db 121 IleGlnAlaAlaPheArgGlyHisValAlaAlaArgGluGluValLysMetLysThrAsp 140

QY 556 AGCTCTTCAAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 588
Db 141 SerLeuGlnAsnGluGluLysGluGluAsnLys 151

RESULT 4
Q9BDQ6 PRELIMINARY; PRT; 143 AA.
ID Q9BDQ6
AC Q9BDQ6;

```

DR SMART; SM00015; IQ; 1.
 DR SMART; SM00394; RIIA; 1.
 DR PROSITE; PS50096; IQ; 1.
 KW Membrane.
 FT DOMAIN 114 143 IQ.
 SQ SEQUENCE 151 AA; 17406 MW; BC99EC310FA0E54A CRC64;

Alignment Scores:
 Pred. No.: 3,62e-58 Length: 151
 Score: 778.00 Matches: 151
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 47.41% Indels: 0
 DB: 1 Gaps: 0

AF334735 (1-954) x SP17_HUMAN (1-151)

QY 136 ATGTCGATTCATTCCTCAACACCCACTACCAAGATTCACAGATTTGGGATCTTCTT 195

Db 1 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeu 20

QY 196 GAAGGGCTGACCGCAGATTCTGAGAGCAACCGGACAAATATACCACTTTTTCAGCA 255

Db 21 GluGlyLeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAla 40

QY 256 GCCTATTTTGGAGCCCTTCTAGAGAAAGAGAAACCAACTTTGATCCAGCAATCG 315

Db 41 AlaTyrPheGluSerLeuGluLeuArgGluGlnProAspAsnIleProAlaPheAla 40

QY 316 GAGAGTAAAGTAGAGACCGCTTCTATACCAATCATGATTCGAGAGCAAGAACCCACT 375

Db 61 GlySerLeuValGluAspArgPheTyrAsnAsnHisAlaPheGluGluGlnPro 80

QY 376 GAGAAAGTGATCTTAAACAAGAGAGTCTCAGATATCTGGGAGGAGAGAGATCA 435

Db 81 GluLysSerAspProLysGlnGluSerGlnIleProGlyLysGluGluGluThrSer 100

QY 436 GTCACATCTTAGACTCTTCTGAGAGAGTAAAGAGAGAGAGAGAGAGAGAGAGAG 495

Db 101 ValThrIleLeuAspSerSerGluGluAspLysGluGluValAlaAlaValLys 120

QY 496 ATCCAAAGCTGCTTCCGGGGACACATAGCCAGAGAGGAGGCAAGAAATGAAACAAT 555

Db 121 IleGlnAlaAlaPheArgGlyHisIleAlaArgGluGluAlaLysMetLysThrAsn 140

QY 556 AGCTCTTCAAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 588

Db 141 SerLeuGlnAsnGluGluLysGluGluAsnLys 151

RESULT 3

SP17_MACFA

ID SP17_MACFA STANDARD; PRT; 151 AA.

AC Q19021;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Sperm surface protein Sp17 (Sperm autoantigenic protein 17).

GN Name=SP17; Synonyms=SP17;

OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

OC Cercopithecinae; Macaca.

OX NCBI_TaxID=9541;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Testis;

RA Lea I.A., O'Rand M.G.;

RT "The immune response to immunization with sperm antigens in the

RT Macaque oviduct."

RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Sperm surface zona pellucida binding protein. Helps to

CC bind spermatozoa to the zona pellucida with high affinity. Might

CC function in binding zona pellucida and carbohydrates (By

DT	01-JUN-2001 (TEMBLrel. 17, Created)
DT	01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT	01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE	Sperm protein 17 (Fragment).
OS	Macaca mulatta (Rhesus macaque).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC	Cercopithecoidea; Macaca.
OX	NCBI_TaxID=9544;
OX	NCBI_TaxID=9544;
RN	1
RP	SEQUENCE FROM N.A.
RC	TISSUE=Splice;
RE	MEDLINE=2280349; PubMed=12393185; DOI=10.1016/S0167-4781(02)00478-5;
RX	Buchli R., De Jong A., Robbins D.L.;
RT	"Genomic organization of an intron-containing sperm protein 17 gene
RT	(Spl7-1) and an intronless pseudogene (Spl7-2) in humans: a new
RT	model";
RL	Biochim. Biophys. Acta 1578:29-42 (2002).
DR	EMBL; AF334809; AAK28124.1; -.
DR	HSSP; P12367; 1R2A.
DR	GO; GO:0008603; F:cAMP-dependent protein kinase regulator act. . . ; IEA.
DR	GO; GO:0007165; P:signal transduction; IEA.
DR	InterPro: IPR000048; IQ_region.
DR	InterPro: IPR003117; RIIA.
DR	Pfam; PF00612; IQ; 1.
DR	Pfam; PF02197; RIIA; 1.
DR	SMART; SM00015; IQ; 1.
DR	SMART; SM00394; RIIA; 1.
DR	PROSITE; PS50096; IQ; 1.
FT	NON_TER 1
FT	NON_TER 143
FT	NON_TER 143
SQ	SEQUENCE 143 AA; 16424 MW; 70B538F7F876B465 CRC64;
Alignment Scores:	
Pred. No.:	2,528-52 Length: 143
Score:	710.00 Matches: 137
Percent Similarity:	97.20% Conservative: 2
Best Local Similarity:	95.80% Mismatches: 4
Query Match:	43.27% Indels: 0
DB:	2 Gaps: 0
AF334735 (1-954) x Q9BDQ6 (1-143)	
QY	154 AACACCCACTACCGAATCCACAGGATTTGGGAATCTTTTGAAGGGCTGACACGCGAG 213
DB	1 AsnThrHisTyArgIleProGlnGlyPheGlyAsnLeuLeuGluGlyLeuThrArgGlu 20
QY	214 ATTCTGAGAGACAACCGACAATATACCACTTTTGCAGCAGCTATTTTGAGAGCCCT 273
DB	21 IleLeuArgGluGlnProAspAsnIleProAlaPheAlaAlaTyPheGluSerLeu 40
QY	274 CTAGAGAAAGAGAGAAACCACTTTCATCCAGCAGATGCGGAGTAGGTAGAGAC 333
DB	41 LeuGluLysArgGluLysThrAsnPheAspProAlaGluTrpGlySerLysValGluAsp 60
QY	334 CGCTTCTATAACAATCATCTCATCTCCAGAGACAAGAACCCACTGTAGAAAAGTGTCTTAA 393
DB	61 ArgPheTyAsnAsnHisAlaPheGluGlnGlyProProGluLysSerAspProLys 80
QY	394 CAAGAAGAGTCTCAGATATCTCGGAAGGAGGAAGACATCAGTCACCATCTTAGACTCT 453
DB	81 GlnGluSerGlnIleProGlyLysGluGluAlaSerValThrIleLeuAspSer 100
QY	454 TCTGAGAGAGATAGAAAAGACAGAGTGTCTGCTGTCAAAATCCAGCTGCTTCCGG 513
DB	101 SerGluGluAspLysGluLysGluValAlaValLysIleGlnAlaAlaPheArg 120
QY	514 GGACACATACCGAGAGAGGAGCAAGAAAATGAAACAAATAGTCTTCAAAATGAGGAA 573
DB	121 GlyHisValAlaArgGluGluValLysLysMetLysThrAspSerLeuGlnAsnGluGlu 140
QY	574 AAAGAGGAA 582

```
QY 442 ATCTTAGACTCTTCTGAGGAGATAGAGAAAAAGAGAGTGTCTGCTGTCATAATCCAA 501
Db 161 IleLeuAspSerSerGluGluAspLysGluLysGluGluValAlaAlaValLysIleGln 180
QY 502 GCTGCTCTCCGGGACACATAGCAGAGAGGAGGCAAGAAATGAACAATAGTCTT 561
Db 181 AlaAlaPheArgGlyHisIleAlaArgGluGluValLysMetLysThrAspSerLeu 200
QY 562 CAAAAT-----GAGGAAAAAGAGGAAAAAAG 588
Db 201 GlnAsnGluLysGluGluLysGluGluAsnLys 212

RESULT 6
Q9TU08 PRELIMINARY; PRT; 147 AA.
AC Q9TU08;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE SP17 protein.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OC NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Jejunal Peyer's patch;
RX MEDLINE=21015023; PubMed=1132159; DOI=10.1007/s002510000253;
RA Tatlow D., Brownlie R., Babluk L.A., Griebel P.;
RT "Differential display analysis of gene expression during the induction
of mucosal immunity."
RL Immunogenetics 52:73-80(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Jejunal Peyer's patch;
RA Tatlow D.D., Brownlie R., Babluk L.A., Griebel P.J.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179926; AAD54394.1; -.
DR HSSP; P12367; 1R2A.
DR GO; GO:0008603; F:GMP-dependent protein kinase regulator act. . . ; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000048; IQ region.
DR InterPro; IPR003117; RIIa.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF02197; RIIa; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00394; RIIa; 1.
DR PROSITE; PS50096; IQ; 1.
SQ SEQUENCE 147 AA; 16903 MW; 4F9CB44103C25B66 CRC64;

Alignment Scores:
Pred. No.: 6,34e-42 Length: 147
Score: 589.00 Matches: 115
Percent Similarity: 86.09% Conservative: 15
Best Local Similarity: 76.16% Mismatches: 17
Query Match: 35.89% Indels: 4
DB: 2 Gaps: 1

AF334735 (1-954) x Q9TU08 (1-147)
QY 136 ATGTGATTCATTCTCCACACCCACTACCGAATTCACAGATTTCGGAATCTTCT 195
Db 1 MetSerIleSerPheSerAsnThrHisTyArgIleProGlnGlyPheGlyAsnLeuLeu 20
QY 196 GAAGGGCTGACCGCGAGATTCTGAGAGAGCAACCGGACAATATACAGCTTTTGAGCA 255
Db 21 GluGlyLeuThrArgGluIleLeuArgGluGlnProAspAsnIleProAlaPheAla 40
QY 256 GCTATTTTGAGAGCTTCTAGAGAAAGAGAGAAACCACTTTGATCCAGCAGATGG 315
Db 41 AlaTyPheGluAsnLeuLeuGluLysArgGluLysThrSerPheAspProAlaGluTrp 60
```

```
QY 316 GGGAGTAAAGTAGAAGACCGCTTCTATAACATCATGCAATTCGAGGAGCAAGACCACT 375
Db 61 GlyAlaLysValAspAspArgPheTyraenAsnHisAlaPheLysGluGlnGluSerPro 80
QY 376 GAGAAAAGTATCTCTAAACAAGAGTCTCAGATATCTGGGAAGAGGAGAGAGACATCA 435
Db 81 GluLysHisGluProGlyLysGluAsnSerGlnThrSerValLysGluLys----- 97
QY 436 GTCCACCATCTAGACTCTTCTGAGGAAGATAAGGAAAAAGAGAGGTTCTGCTGTCAAA 495
Db 98 ---ThrAlaLeuGluSerProGluGlyAspLysAspMetGluGluAsnAlaLeuLys 116
QY 496 ATCCAAGTCTCTCCGGGACACATAGCAGAGAGGCGCAAAAGAAAATGAACAAT 555
Db 117 IleGlnAlaAlaPheArgGlyHisLeuAlaArgGluGluValLysMetLysSerIle 136
QY 556 AGTCTTCAAAATGAGGAAAAAGAGGAAAAAAG 598
Db 137 AspLeuGluGluLysThrGluGluAsnLys 147

RESULT 7
SP17_RABIT
ID SP17_RABIT STANDARD; PRT; 146 AA.
AC P36425;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Sperm surface protein Sp17 (Sperm autoantigenic protein 17).
GN Name=SPAI7; Synonyms=SP17;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OC NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 113-126.
RC STRAIN=New Zealand white; TISSUE=Testis;
RX MEDLINE=95046885; PubMed=7525387; DOI=10.1006/dbio.1994.1285;
RA Richardson R.T., Yamasaki N., O'Rand M.G.;
RT "Sequence of a rabbit sperm zona pellucida binding protein and
localization during the acrosome reaction."
RL Dev. Biol. 165:688-701(1994).
CC -!- FUNCTION: Sperm surface zona pellucida with high affinity. Might
bind spermatozoa to the zona pellucida and carbohydrates.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Membrane-associated (Potential).
CC -!- TISSUE SPECIFICITY: Testis- and sperm-specific.
CC -!- PTM: The N-terminus is blocked.
CC -!- SIMILARITY: Contains 1 IQ domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z20655; CAA79674.1; -.
CC PIR; I46506; I46506.
CC HSSP; P12367; 1R2A.
CC InterPro; IPR000048; IQ region.
CC InterPro; IPR003117; RIIa.
CC Pfam; PF00612; IQ; 1.
CC Pfam; PF02197; RIIa; 1.
CC SMART; SM00015; IQ; 1.
CC SMART; SM00394; RIIa; 1.
CC PROSITE; PS50096; IQ; 1.
KW Direct protein sequencing; Membrane.
FT DOMAIN 110..139 IQ.
FT SEQUENCE 146 AA; 16891 MW; 3FF302B883D9566F CRC64;
SQ
```

Alignment Scores:
Pred. No.: 3.32e-40 Length: 146
Score: 569.00 Matches: 109
Percent Similarity: 86.67% Conservative: 21
Best Local Similarity: 72.67% Mismatches: 16
Query Match: 34.67% Indels: 4
DB: 1 Gaps: 3

AF334735 (1-954) x SP17_RABIT (1-146)

QY 136 ATGTCGATCCATTCTCCACACCCTACCGAATTCACAGGATTTGGGAATCTTCTT 195
Db 1 MetSerIleProPheSerAsnThrHisTyArgIleProGInglyPheGlyAsnLeu 20
QY 196 GAAGGGCTCACGCCGAGATTCTGAGAGAGCAACCGACAATATACCAAGCTTTTGCGACA 255
Db 21 GluGlyeuThrArgGluileLeuArgGluInProAspAsnIleProAlaPheAla 40
QY 256 GCCTATTTGAGAGCCTTCTAGAGAAAAGAGAGAAACCAACTTTGTATCCAGCAGAATGG 315
Db 41 AlaTy-PheGluAsnLeuLeuArgGluLysThyAsnPhesPheProAlaGluTrp 60
QY 316 GGGAGTAAGGTAGAACCGCTCTTATAACATCATGCATTCGAGGAGCAAGACCCT 375
Db 61 GlyAlaLysValGluAspArgPheTyAsnAsnHISalAPheGlnGluHIEglu---Ser 79
QY 376 GAGAAAAGTGATCTCTAAACAAGAGAGTCTCAGATATCTGGGAAGGAGGAGAGACATCA 435
Db 80 GluLysCyseGlu-----GlnGluLysSerGlnSerValThr-----GluGluGluThrPro 97
QY 436 GTCCACCATCTTAGACTCTTCTGAGGAAGATAAGGAAAAAGAGAGGTTGCTGCTGTCAA 495
Db 98 ValLeuThrIleasp---SerGluAspAspLysAspLysGluGluMetAlaAlaLys 116
QY 496 ATCCAAGTCCTTCGGGGACACATAGCCAGAGAGGCGCAAGAAAATGAAACAAT 555
Db 117 IleGlnAlaAlaPheArgGlyHisValAlaArgGluAspValLysIleArgThrAsn 136
QY 556 AGTCTTCAAAATGAGGAAAAAGGAAAC 585
Db 137 LysAlaGluGluThrGluGluAsn 146

RESULT 8
Q921K2 PRELIMINARY; PRT; 148 AA.
ID Q921K2 AC Q921K2
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-OCT-2003 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sp17 protein...
GN Names=sp17;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Frayne J., Jury J.A., Hall L.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ131888; CAAL0524.1; -;
DR HSSP; P12367; 1R2A.
DR GO; GO:0008603; F:CAMP-dependent protein kinase regulator act. . . ; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR003117; RIIfa.
DR DR Pfam; PF00612; IQ; 1.
DR DR SMART; SM00015; IQ; 1.
DR DR SMART; SM00394; RIIfa; 1.
DR DR PROSITE; PS00036; IQ; 1.
SQ SEQUENCE 148 AA; 17096 MW; 5E0BC9D4C8FD585 CRC64;

Alignment Scores:
Pred. No.: 1.46e-39 Length: 148
Score: 561.50 Matches: 109
Percent Similarity: 86.67% Conservative: 21
Best Local Similarity: 72.67% Mismatches: 17
Query Match: 34.22% Indels: 3
DB: 2 Gaps: 2

AF334735 (1-954) x Q921K2 (1-148)

QY 136 ATGTCGATCCATTCTCCACACCCTACCGAATTCACAGGATTTGGGAATCTTCTT 195
Db 1 MetSerIleProPheSerAsnThrHisTyArgIleProGInglyPheGlyAsnLeu 20
QY 196 GAAGGGCTCACGCCGAGATTCTGAGAGAGCAACCGACAATATACCAAGCTTTTGCGACA 255
Db 21 GluGlyeuThrArgGluileLeuArgGluInProAspAsnIleProAlaPheAla 40
QY 256 GCCTATTTGAGAGCCTTCTAGAGAAAAGAGAGAAACCAACTTTGTATCCAGCAGAATGG 315
Db 41 AlaTy-PheGluAsnLeuLeuArgGluLysThyAsnPhesPheProAlaGluTrp 60
QY 316 GGGAGTAAGGTAGAACCGCTCTTATAACATCATGCATTCGAGGAGCAAGACCCT 375
Db 61 GlyAlaLysValGluAspArgPheTyAsnAsnHISalAPheLysAspProGluGlnAla 80
QY 376 GAGAAAAGTGATCTCTAAACAAGAGAGTCTCAGATATCTGGGAAGGAGGAGAGACATCA 435
Db 81 GluLysCyseGlu-----GlnGluLysAlaLysAlaSerGlyArg---GluGluThrPro 97
QY 436 GTCCACCATCTTAGACTCTTCTGAGGAAGATAAGGAAAAAGAGAGGTTGCTGCTGTCAA 495
Db 98 ValThrProPheGluGluSerThrGluGluGluArgGluGlnGluAlaAlaValLys 117
QY 496 ATCCAAGTCCTTCGGGGACACATAGCCAGAGAGGCGCAAGAAAATGAAACAAT 555
Db 118 IleGlnSerAlaPheArgGlyHisValAlaArgGluGluValLysMetLysSerAsp 137
QY 556 AGTCTTCAAAATGAGGAAAAAGGAAAC 585
Db 138 LysSerGluAsnValLysGluGluAsn 147

RESULT 9
SP17 MOUSE
ID _SP17_MOUSE STANDARD; PRT; 149 AA.
AC Q62252;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Sperm surface protein Sp17 (Sperm autoantigenic protein 17).
GN Name=Sp17; Synonyms=Sp17;
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Testis;
RX MEDLINE=96039129; PubMed=7578692;
RA Kong M., Richardson R.T., Widgren B.E., O'Rand M.G.;
RT "Sequence and localization of the mouse sperm autoantigenic protein,
RT Sp17";
RL Biol. Reprod. 53:579-590(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Hopkins R.F., Zeeberg B., Buettow K.H., Schaefer C.P., Bhat N.K.,
RA Stapleton M., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Murzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: Sperm surface zona pellucida binding protein. Helps to
 CC bind spermatozoa to the zona pellucida with high affinity. Might
 CC function in binding zona pellucida and carbohydrates (By
 CC similarity).
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Membrane-associated (Potential).
 CC -!- TISSUE SPECIFICITY: Testis- and sperm-specific.
 CC -!- SIMILARITY: Contains 1 IQ domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z46299; CAAB6455.1; -.
 DR EMBL; BC059727; AAH59727.1; -.
 DR PIR; S49527; S49527.
 DR HSSP; P12367; 1R2A.
 DR MGD; MGI:1333778; Spal7.
 DR InterPro; IPR000048; IQ_region.
 DR InterPro; IPR003117; RIIa.
 DR Pfam; PF00612; IQ; 1.
 DR SMART; SM00197; RIIa; 1.
 DR SMART; SM00394; RIIa; 1.
 DR PROSITE; PS50096; IQ; 1.
 KW Membrane.
 FT DOMAIN 112 141 IQ.
 SQ SEQUENCE 149 AA; 17296 MW; C7E05D111D6AF0DC CRC64;
 Alignment Scores:
 Pred. No.: 2,93e-39 Length: 149
 Score: 558.00 Matches: 107
 Percent Similarity: 86.49% Conservative: 21
 Best Local Similarity: 72.30% Mismatches: 18
 Query Match: 34.00% Indels: 2
 DB: 1 Gaps: 1
 AP334735 (1-954) x SP17_MOUSE (1-149)
 QY 136 ATGTCGATTCATTTCTCCACACCCCACTACCGAATCCCAAGGATTTGGGAATCTTCTT 195
 Db |||||
 QY 196 GAAGGCTCACCGCAGATCTTGAGAGACCAACCGGACAATATACCACTTTTGCAGCA 255
 Db |||||
 QY 21 GluGlyLeuThrArgGluIleuArgGluGlnProAspAsnIleProAlaPheAlaLa 40
 QY 256 GCCTATTTTGAGAGCTCTTAGAGAAAGAGAGAAACCACTTTGATCCACGACGATCG 315
 Db |||||
 QY 41 AlaTyrPheGluAsnLeuLeuGluIleGluGlySerPheAspProAlaGluIle 60
 QY 316 GGGAGTAAGGTAGAAGACCGCTTCTATAACAATCATGCAATCATGAGGAGCAAGAACCCACT 375
 Db |||||
 QY 61 GlyAlaIleValGluAspArgPheTyrAsnAsnHisAlaPheIleGluGlnVal 80
 QY 376 GAGAAAGTGATCTCTAAACAGAGAGTCTCAGATATCTGGGAGGAGGAGAGACATCA 435
 Db |||||

Db 81 GluIysCysGlu-----GlnGluLeuAlaIysSerSerGlyArgGluGluThrProVal 98
 QY 436 GTCACCATCTTAGACTCTTCTGAGGAAGATAGGAAAAAGAGAGGTCTGCTGTCACAA 495
 Db ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
 QY 99 ThrProPheGluGluSerThrGluGluGluArgGluGluGluGluAlaAlaLeuIys 118
 QY 496 ATCCAAAGTCCTTCGCGGACACATAGCCAGAGAGAGGCGCAAGAAATGAAAAAAT 555
 Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
 QY 119 IleGlnSerLeuPheArgGlyHisValAlaArgGluGluValIysLysMetLysSerAsp 138
 QY 556 AGTCTTCAATGAGAAAAAGAG 579
 Db ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
 QY 139 LysAenGluAsnLeuIysGluGlu 146
 Db
 RESULT 10
 SP17_MACEU
 ID SP17_MACEU STANDARD; PRT; 153 AA.
 AC O62770;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Sperm surface protein Sp17 (Sperm autoantigenic protein 17).
 GN Name=SPAL7; Synonyms=SP17;
 OS Macropus eugenii (Tamar wallaby).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
 CC NCBI TaxID=9315;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Wen Y., O'Rand M.G.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Sperm surface zona pellucida binding protein. Helps to
 CC bind spermatozoa to the zona pellucida with high affinity. Might
 CC function in binding zona pellucida and carbohydrates (By
 CC similarity).
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Membrane-associated (Potential).
 CC -!- TISSUE SPECIFICITY: Testis- and sperm-specific (By similarity).
 CC -!- SIMILARITY: Contains 1 IQ domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF054289; AAC08024.1; -.
 DR HSSP; P12367; 1R2A.
 DR InterPro; IPR000048; IQ_region.
 DR InterPro; IPR003117; RIIa.
 DR Pfam; PF00612; IQ; 1.
 DR Pfam; PF02197; RIIa; 1.
 DR SMART; SM00015; IQ; 1.
 DR SMART; SM00394; RIIa; 1.
 DR PROSITE; PS50096; IQ; 1.
 KW Membrane.
 FT DOMAIN 122 151 IQ.
 SQ SEQUENCE 153 AA; 17327 MW; AD3A7C6B4D4B1B65 CRC64;
 Alignment Scores:
 Pred. No.: 4.69e-38 Length: 153
 Score: 544.00 Matches: 104
 Percent Similarity: 81.05% Conservative: 20
 Best Local Similarity: 67.97% Mismatches: 28
 Query Match: 33.15% Indels: 8
 DB: 1 Gaps: 1
 AP334735 (1-954) x SP17_MACEU (1-153)
 QY 136 ATGTCGATTCATTTCTCCACACCCCACTACCGAATCCCAAGGATTTGGGAATCTTCTT 195

```

Db      1 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeuLeu 20
QY      196 GAAGGCTGACACCGGAGATTCTGAGAGACCAACCGGACATATATACCAAGCTTTTCAGCA 255
Db      21 GluGlyLeuThrArgGluIleLeuArgGluGlnProAspAsnIleProAlaPheAlaAla 40
QY      256 GCCTATTGAGAGCTCTGAGAGAAAGAGAGAAACCAACTTTGATCCAGCAATGG 315
Db      41 AlaTyrPheGluAsnLeuGluLysArgGluLysThrAsnPheAspProAlaGluTrp 60
QY      316 GGGAGTAAAGTAAAGACCGCTTCTATACCAATCATGCTATTCGATTCGAG 360
Db      61 GlyAlaLysIleAspAspArgPheTyrAsnAsnHisAlaPheLysValProSerGlyAla 80
QY      361 -----GAGCAAGAACCCCTCAGAAAGTGTCTTAAACAAGAGAGTCTCAGATA 411
Db      81 ThrGluSerLysGluAlaProProGluLysSerGluProGluLysGluThrProGlnGlu 100
QY      412 TCTGGAGGAGGAGAGACATCAGTCACCATCTTAGACTCTCTGAGAGAGATAAGGAA 471
Db      101 ValValLysGluGlnGluThrGlnValSerPheValGluGluValSerThrAspAspGlu 120
QY      472 AAAGAAGAGGTTGCTGCTGCTCAAAATCCAAAGTGTCTCCGGGGACACATAGCCAGAGAG 531
Db      121 GluAlaAlaAlaAlaAlaValLysIleGlnAlaAlaPheArgGlyHisLysAlaArgLys 140
QY      532 GAGGCAAGAAATGAAACAATAGTCTTCAAAATGAG 570
Db      141 GluValLysIleMetLysGluSerSerIleGluGluGln 153

RESULT 11
SP17_MONDO
ID SP17 MONDO STANDARD; PRT: 179 AA.
AC O62771;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Sperm surface protein Sp17 (Sperm autoantigenic protein 17).
GN Name=SPAL7; Synonyms=SP17;
OS Monodelphis domestica (Short-tailed grey opossum).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Monodelphis.
OX NCBI_TaxID=13616;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Wen Y., O'Rand M.G.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Sperm surface zona pellucida binding protein. Helps to
CC bind spermatozoa to the zona pellucida with high affinity. Might
CC function in binding zona pellucida and carbohydrates (By
CC similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Membrane-associated (Potential).
CC -!- TISSUE SPECIFICITY: Testis, and sperm-specific (By similarity).
CC -!- SIMILARITY: Contains 1 IQ domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF054290; AAC08025.1; -.
CC HSP: P12367; I82A.
CC InterPro: IPR000048; IQ_region.
CC DR InterPro; IPR003117; RIIA.
CC DR Pfam; PF00612; IQ; 1.
CC DR Pfam; PF02197; RIIA; 1.
CC SMART; SM00015; IQ; 1.

```

```

DR SMART; SM00394; RIIA; 1.
DR PROSITE; PS50096; IQ; 1.
KW Membrane.
FT DOMAIN 124 132 Poly-Glu.
FT DOMAIN 143 172 IQ.
SQ SEQUENCE 179 AA; 20864 MW; CADDCL13CEC66A00E CRC64;

Alignment Scores:
Pred. No.: 1.75e-34 Length: 179
Score: 502.50 Matches: 105
Percent Similarity: 69.44% Conservative: 20
Best Local Similarity: 58.33% Mismatches: 22
Query Match: 30.62% Indels: 33
Db: 1 Gaps: 4

AF334735 (1-954) x SP17_MONDO (1-179)
QY 136 ATGTCGATTCCTCATTCCTCAACACCCACCTACCAAGGATTTCCACAGGATTTGGGAATCTTCTT 195
Db 1 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeu 20
QY 196 GAAGGCTGACACCGGAGATTCTGAGAGACCAACCGGACATATATACCAAGCTTTTCAGCA 255
Db 21 GluGlyLeuThrArgGluIleLeuArgGluGlnProAspAsnIleProAlaPheAlaAla 40
QY 256 GCCTATTGAGAGCTCTGAGAGAAAGAGAGAAACCAACTTTGATCCAGCAATGG 315
Db 41 AlaTyrPheGluAsnLeuGluLysArgGluLysThrAsnPheAspProAlaGluTrp 60
QY 316 GGGAGTAAAGTAAAGACCGCTTCTATACCAATCATGCTATTCGAGAGAGAGAACACCACT 375
Db 61 GlyAlaLysValAspAspArgPheTyrAsnAsnHisAlaPheLysAsp---GluProPro 79
QY 376 GAGAAAGTGTAT-----CCTAAACAAGAGAG 402
Db 80 GluLysSerGluThrGlnLysIleGlnProGluLysValAlaIleGluLysGluThrMet 99
QY 403 TCTCAGATATCTGGAGAGGAGAGACATCAGTCACCATCTTAGACTCTTCGAGGAA 462
Db 100 ProGlnGluThrValLysGluLysGluThrGlnValSerPheValGluGluProThrGlu 119
QY 463 GATAAGGAAAGAGAGGTT----- 483
Db 120 GluProGlnLysGluGluGluGluGluGluGluGluGluGluGluGluGluGluVal 139
QY 484 -----GCTGCTGTCAAAATCAAGTCCTTCGGGGGACACATAGCC 525
Db 140 ArgGluGlyMetGlnAspAlaAlaValLysIleGlnAlaValPheArgGlyHisLysThr 159
QY 526 AGAGAGAGGCAAGAGAAATCAAAACAATAGTCTTCAAAATGAGGAGGAGGAGGAGGAGG 585
Db 160 ArgLysGluTyrLeuLysLysArgAspSerThr-----AspGluThrAlaAspGluAsn 177

RESULT 12
Q6SJ92
ID Q6SJ92 PRELIMINARY; PRT: 141 AA.
AC Q6SJ92;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Sperm autoantigenic protein 17.
GN Name=SPAL7;
OS Ornithorhynchus anatinus (Duckbill platypus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Monotremata; Ornithorhynchidae; Ornithorhynchus.
OX NCBI_TaxID=9258;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=21308450; PubMed=1415432; DOI=10.1042/0264-6021.3570025;
RA Wen Y., Richardson R.T., Widgeon E.E., O'Rand M.G.;
RT "Characterization of Sp17: a ubiquitous three domain protein that
RT binds heparin."

```

```

RL Biochem. J. 357:25-31 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.; to the EMBL/GenBank/DBJ databases.
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC075533; AAH75533.1; -.
DR GO; GO:0008603; F:camp-dependent protein kinase regulator act. . . IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000048; IQ region.
DR InterPro; IPR003117; RIIa.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF02197; RIIa; 1.
DR SMART; SM00015; IQ; 2.
DR SMART; SM00394; RIIa; 1.
DR PROSITE; PS50096; IQ; 2.
SQ SEQUENCE 495 AA, 56219 MW, 4BB57410B790912A CRC64;

Alignment Scores:
Pred. No.: 1,52e-22 Length: 495
Score: 364.00 Matches: 80
Percent Similarity: 58.42% Conservative: 31
Best Local Similarity: 42.11% Mismatches: 35
Query Match: 22.18% Indels: 44
DB: 2 Gaps: 5

AF334735 (1-954) x Q6DIJ9 (1-495)
QY 136 ATGTCGATTCATCTCCACACCCACTACCGAATTCACAGGATTTGGGATCTTCTT 195
DB 1 MetSerileProPheSerAsnThrHisTyrArgIleProArgGlyPheAlaAsnLeu 20
QY 196 GAAGGCTGACGCGGAGATTCTGAGAGAGCAACCGCAATATATACACGCTTTTCGAGCA 255
DB 21 GluGlyLeuThrArgGluLeuValLeuArgHisGlnProLysAspIleProLeuPheGlyAla 40
QY 256 GCCTATTTTGAGAGCTTTCTAGAGAAAAGAGAAAACCAACTTTTGATCCAGCAATGG 315
DB 41 LysTyrPheSerGlnLeuGlnArgGlnAspThrAspPheAspProAlaGlnTyr 60
QY 316 GCGAGTAGGTAGAGACCGCTTCTATACATCATCATGATTCGAGGAGCAAGACCACT 375
DB 61 GlyAlaAlaLeuGluAspArgPheTyrAsnAsnTyrAspPheGlnHisThrGluLeuLys 80
QY 376 GAGAAAAGTGATCTCTAAACAAGAGAGTCTCAGATATCTGGGAGAGAGAGAGACATCA 435
DB 81 LysTyrThr-----SerGluProSerGluLeuLys-----LysSerLysArgLysSer 96
QY 436 GTCCACATCTTAGACTCTTCTGAG-----GTCTCTGAG-----GAAGAT 465
DB 436 GTCCACATCTTAGACTCTTCTGAG-----GTCTCTGAG-----GAAGAT 465

NCBI_TaxID=8364;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.; to the EMBL/GenBank/DBJ databases.
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC075533; AAH75533.1; -.
DR GO; GO:0008603; F:camp-dependent protein kinase regulator act. . . IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000048; IQ region.
DR InterPro; IPR003117; RIIa.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF02197; RIIa; 1.
DR SMART; SM00015; IQ; 2.
DR SMART; SM00394; RIIa; 1.
DR PROSITE; PS50096; IQ; 2.
SQ SEQUENCE 141 AA, 15752 MW, 98E3282E0314D0B0 CRC64;

Alignment Scores:
Pred. No.: 1,51e-33 Length: 141
Score: 491.50 Matches: 102
Percent Similarity: 77.48% Conservative: 15
Best Local Similarity: 67.55% Mismatches: 23
Query Match: 29.95% Indels: 11
DB: 2 Gaps: 4

AF334735 (1-954) x Q6SU92 (1-141)
QY 136 ATGTCGATTCATCTCCACACCCACTACCGAATTCACAGGATTTGGGATCTTCTT 195
DB 1 MetSerileProPheSerAsnThrGlnTyrArgIleProGlnGlyPheGlyAsnValLeu 20
QY 196 GAAGGCTGACGCGGAGATTCTGAGAGAGCAACCGCAATATATACACGCTTTTCGAGCA 255
DB 21 GluGlyLeuThrArgGluLeuLysGluGlnProAspAsnIleProAlaPheAlaAla 40
QY 256 GCCTATTTTGAGAGCTTTCTAGAGAAAAGAGAAAACCAACTTTTGATCCAGCAATGG 315
DB 41 AlaTyrPheGluAlaLeuGluLysArgGluLysSerAsnPheAspProAlaGlnTyr 60
QY 316 GCGAGTAGGTAGAGACCGCTTCTATACATCATCATGATTCGAGGAGCAAGACCACT 375
DB 61 GlyAlaLysAlaAspAspArgPheTyrAsnAsnGlnAlaPheLysGluAlaAlaThr 80
QY 376 GAGAAAAGTGATCTCTAAACAAGAGAGTCTCAGATATCTGGGAGAGAGAGACATCA 435
DB 81 Glu-----ProGluLysValLysSerAlaValSerValLys---GluGluAlaLys 96
QY 436 GTCCACATCTTAGACTCTTCTGAGGAGATAGGAAAAA---GAAGAGTTCTGCTGTC 492
DB 97 ValThrProLeuGluProGlnAlaAspValAsnArgGlyGluAsnGluAlaAlaThr 116
QY 493 AAAATCCAGCTGCTTCCTCGGAGACATGATGCGAGAGAGGAGGAGGAGGAGGAGGAGGAG 552
DB 117 LysileGlnAlaAlaPheArgGlyHisLeuAlaArgGluGluLeuLysLysMetLysAla 136
QY 553 AATAGTCTTCAAAATGAGGAAAAAGAGGAAAAAC 585
DB 137 -----GluAlaGluGluAsn 141

RESULT 13
Q6DIJ9 PRELIMINARY; PRT; 495 AA.
AC Q6DIJ9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Spal7-prov protein.
GN Name=spal7-prov;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.

```

```

Db      97 PheSerMetLeuAspAsnSerAspThrTyrProLeuSerProSerThrSerLeuGluGln 116
      :::::|||||:::|||||:::
QY      466 AAGGAA-----AAAGAAGAGTTGCTGCTGTCATAAATCAAGCTGCC 507
      |||||
Db      117 LysGluIleGlnThrAspMetArgLysMetGlnAlaAlaThrThrIleGlnAlaAla 136
      |||||
QY      508 TTCGGGGGACACATAGCCAGA----- 528
      |||||:::|
Db      137 PheArgGlyTyrAsnValArgHisLysValLysIleHisGlnGluSerPheValAspAsn 156
      |||||:::|
QY      529 -----GAGGAGGCAAAAGAAATGAAAAACAAT 555
      :::::|||||:::|||||:::
Db      157 TyrMetLysProGluProGluHisGluAspThrAspGlnAlaLysGluThrLysCysAsp 176
      :::::|||||:::|||||:::
QY      556 AGCTTTCAAAATGAGGAAAGAGAAAC 585
      :::::|||||:::|
Db      177 LeuMetGluAsnGluAspCysGluPheAsn 186

RESULT 14
Q20199 ID Q20199 PRELIMINARY; PRT; 189 AA.
AC Q20199;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein F39H12.3.
GN Name=F39H12.3; ORFNames=F39H12.3;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Minx M., Wohldmann P.;
RT "The sequence of C. elegans cosmid F39H12.";
RN Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; U42847; AA883605.1; -.
DR PIR; T29159; T29159.
DR WormBase; WBGene00018214; F39H12.3.
DR WormPep; F39H12.3; CB07210.
DR GO; GO:0008603; F.cAMP-dependent protein kinase regulator act. . .; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000048; IQ region.
DR InterPro; IPR003117; RIIa.
DR Pfam; PF00612; IQ; 3.
DR Pfam; PF02197; RIIa; 1.
DR SMART; SM00015; IQ; 2.
DR SMART; SM00394; RIIa; 1.

```

```

DR PROSITE; PS50096; IQ; 2.
KW Hypothetical protein.
SQ SEQUENCE 189 AA; 21625 MW; 0C640476338DE493 CRC64;

Alignment Scores:
Pred. No.: 0.00825 Length: 189
Score: 134.00 Matches: 47
Percent Similarity: 39.25% Conservative: 26
Best Local Similarity: 25.27% Mismatches: 44
Query Match: 8.17% Indels: 69
DB: 2 Gaps: 7

AF334735 (1-954) x Q20199 (1-189)
QY 169 ATTCACACAGATTGGGAATCTTCAAGGGCTGACACGGAGATTCTCAGAGAGCAA 228
      ::|||
Db 9 ValProHisAspLeuArgProIleLeuAlaLeuAlaArgGluValLeuArgSerGln 28
      ::|||
QY 229 CCGGACAATATACCAGCTTTTGCAGCAGCCCTATTGAGAGCCTTCTAGAG---AAAAGA 285
      |||
Db 29 ProSerAspValAlaGluPheGlyHisMetPhePheAspGluTyrLeuLysHisArgArg 48
      |||
QY 286 GAGAAACCAACTTT-----GATCCAGCAGAAATGGGGAGTAAGGTAGAGAAGCGCTTC 339
      |||
Db 49 GluAsnArgAsnIleLeuLysAspProAla----- 58
      |||
QY 340 TATAACAATCATGATTCGAGGAGCAGAACCACTGAGAAAGTGTATCCTTAACAAGAA 399
      |||
Db 59 -----AlaTyrGlu-----ValPheArgAlaAspLeuGlnLysLys 70
      |||
QY 400 GAGTCTCAGATATCTGGGAAGGAGAGACATCAGTCACCATCTTAGACTCTTCTTCGAG 459
      ::|||
Db 71 PheAlaGluVal-----Glu 75
      |||
QY 460 GAAGATAAGGAAAGAAAGAGAGTTGCTGCTCAAAATCCAAGTGCCTCCGGGACAC 519
      ::|||
Db 76 ArgProAlaSerProMetAspThrAlaAlaThrLysIleGlnAlaAlaPheLysGlyHis 95
      |||
QY 520 ATAGCCAGAGAGGAGGCAAGAAATG-----AAA 549
      ::|||
Db 96 LeuValArgAlaHisProGluLysTyrGlyMetSerThrArgThrSerSerSerGluLys 115
      |||
QY 550 ACAATAGTCTTCAAAATGAGAAAGAGAGAAA----- 583
      |||
Db 116 LeuAspSerAlaAsnAsn-LysLysAspGlnLysArgHisSerValGlyGlyTyrThrI 135
      |||
QY 584 -----ACAAGTGAAGACACTGGTTT 603
      |||
Db 135 eAspValAspThrProGluAspArgAlaAlaThrLysIleGlnSerGluIleArgGlyPh 155
      |||
QY 604 TACTCCAGGAAACAT 619
      |||
Db 155 eLeuThrArgLysHis 160

RESULT 15
Q6NW93 ID Q6NW93 PRELIMINARY; PRT; 422 AA.
AC Q6NW93;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein zgc:85886.
GN ORFNames=zgc:85886;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

```

```

Qy      421 GAGGAGAG 429
Db      |||||:::
        94 GluGluAsp 96

RESULT 16
Q25114 PRELIMINARY; PRT; 368 AA..
ID Q25114 AC Q25114
AC Q25114; PRELIMINARY; PRT; 368 AA..
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Regulatory subunit of cAMP-dependent histone kinase.
OS Hemitecnrotus pulcherimus (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euclinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Hemitecnrotus.
OC NCBI_TaxID=7650;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97210759; PubMed=9057823;
RA Hoshino K., Nomura K., Suzuki N.;
RT "Cyclic-AMP-dependent activation of an inter-phylum hybrid histone-
RT kinase complex reconstituted from sea urchin sperm-regulatory subunits
RT and bovine heart catalytic subunits."
RL Eur. J. Biochem. 243:612-623(1997).
DR EMBL; D83379; BAAL1899.1; -.
DR HSP; P12369; ICK4.
DR GO; GO:0005952; C:cAMP-dependent protein kinase complex; IEA.
DR GO; GO:0008603; F:cAMP-dependent protein kinase regulator act. . ; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR002373; CAMP_kin.
DR InterPro; IPR000595; CNMP_binding.
DR InterPro; IPR003117; Riia.
DR Pfam; PF00027; CNMP_binding_2.
DR Pfam; PF02197; Riia_1.
DR PRINTS; PR00103; CAMPKINASE.
DR SMART; SM00100; GNMP_2.
DR SMART; SM00394; Riia_1.
DR PROSITE; PS00888; CNMP_BINDING_1; 1.
DR PROSITE; PS00889; CNMP_BINDING_2; 2.
DR PROSITE; PS00442; CNMP_BINDING_3; 2.
KW Kinase.
SQ SEQUENCE 368 AA; 41679 MW; DF596699E3579186 CRC64;

Alignment Scores:
Pred. No.: 0.0847 Length: 368
Score: 122.50 Matches: 49
Percent Similarity: 47.05% Conservative: 40
Best Local Similarity: 25.93% Mismatches: 67
Query Match: 7.46% Indels: 33
DB: 2 Gaps: 7

AF334735 (1-954) x Q25114 (1-368)
Qy 160 CACTACCGAATTCACAAGGATTGGGAATCTTCTTGAAAGGCTGCACACGCCGAGATTCTG 3166
Db : :::::::::::::::::::::
2 AsnPhGluileProGluGlyLeuThrAspLeuLeuGlnValAlaValLeu 21
Qy 220 AAGAGGCAACCGCACAAATATACCAGCTTTTCACACACGCTTATTTTGAGACCTTCTAGAG 41
Db : :::::::::::::::::::::
22 ArgGluLysProSerAspLeuValGluPheAlaSerSerTyrPheAlaLysLeuGlnGlu 41
Qy 280 -----AAAAGAGAGAAAAACCAACTTTGTATCCGACAGAAATGGGG 318
Db : :::::::::::::::::::::
42 AsnAsnIleSerLeuGlyGlyArgGlyValThrPheThrAlaProGluAspAlaGlu 61
Qy 319 AGTAGGTTAGAAGACCGC-----TTCTATTAACAATCATGATTCGAGGAGCAA 3666
Db : :::::::::::::::::::::
62 SerAspIleAspAspGluProGluLeuProLysAsnArgTyrAlaArgLysSer 81

```


QY 415 GGAAGGAGGAGGAGACATCAGTCACCATCTTGTAGACTCTTCTGAGGAGAGATAAGGAAAA 474
 Db 103 ThrAspThrAspGluAspValThrArgThrGluThrSerAspLysThrThrGlnPhe 122

QY 475 GAAGAGGTTGCTGCTGTC 492
 Db 123 ProSerValTyAlaVal 128

RESULT 20
 O75952 PRELIMINARY; PRT; 493 AA.
 AC O75952;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Fibrous sheathin II.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testes;
 RX MEDLINE=21678993; PubMed=11820818; DOI=10.1006/dbio.2001.0527;
 RA Naaby-Hansen S., Mandal A., Wolkowicz M.J., Sen B., Westbrook V.A.,
 RA Shetty J., Coonrod S.A., Klotz K.L., Kim Y.H., Bush L.A.,
 RA Flickinger C.J., Herr J.C.;
 RT "CABYR, a novel calcium-binding tyrosine phosphorylation-regulated
 RT fibrous sheath protein involved in capacitation.";
 RL Dev. Biol. 242:236-254(2002).
 DR EMBL; AF088868; AAC35373.1; -.
 DR GenBank; HGNC:15569; CABYR.
 DR GO; GO:0008603; P:CaMP-dependent protein kinase regulator act. . . ; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR003117; RIIA.
 DR Pfam; PF02197; RIIA; 1.
 DR SMART; SM00394; RIIA; 1.
 SQ SEQUENCE 493 AA; 52788 MW; 0911CDA626325C27 CRC64;

Alignment Scores:
 Pred. No.: 0.106 Length: 493
 Score: 121.50 Matches: 35
 Percent Similarity: 46.03% Conservative: 23
 Best Local Similarity: 27.78% Mismatches: 43
 Query Match: 7.40% Indels: 25
 DB: 2 Gaps: 3

AF334735 (1-954) x O75952 (1-493)

QY 169 ATTCCACAGGATTGGGAACTCTTGAAGGCTGACACGCGAGATCTTGAGAGCAA 228
 Db 10 ValProTyrGlyLeuLysThrLeuLeuGluGlyIleSerArgAlaValLeuLysThrAsn 29

QY 229 CCGGACAAATACCAAGCTTTTCAGCAGCAGCTATTTCAGAGCCTCTTAGAGAAAAGAGAG 288
 Db 30 ProSerAsnIleAsnGlnPheAlaAlaTyrPheGlnGluLeuThrMetTyrArgGly 49

QY 289 AAAACCAATTCATCCAGCAGATGGGGGAGTAAGTATAGACCGCTTCTATTAACAAAT 348
 Db 50 AsnThrThrMetAsp-----IleLysAspLeuValLysGlnPhe 62

QY 349 CATGCATTCGAGGACGAACACCTGAGAAAGTATGATCTTAACAGAA----- 399
 Db 63 HisGlnIleLysValGluLysTrpSerGluGlyThrThrProGlnLysLysLeuGluCys 82

QY 400 -----GAGTCTCAGATA-----TCT 414
 Db 83 LeuLysGluProGlyLysThrSerValGluSerLysValProThrGlnMetGluLysSer 102

QY 415 GGAAGGAGGAGAGACATCAGTCACCATCTTGTAGACTCTTCTGAGGAGAGATAAGGAAAA 474
 Db 103 ThrAspThrAspGluAspValThrArgThrGluThrSerAspLysThrThrGlnPhe 122

QY 475 GAAGAGGTTGCTGCTGTC 492
 Db 123 ProSerValTyAlaVal 128

RESULT 21
 Q9GM47 PRELIMINARY; PRT; 596 AA.
 AC Q9GM47;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
 OC Cercopitheidae; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain parietal lobe;
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB049839; BAB16725.2; -.
 DR GO; GO:0003723; P:RNA binding; IEA.
 DR GO; GO:0006412; P:protein biosynthesis; IEA.
 DR InterPro; IPR008938; ARM.
 DR InterPro; IPR003890; IF_eIF4G.
 DR InterPro; IPR007193; Upf2.
 DR Pfam; PF02854; MIF4G; 2.
 DR Pfam; PF04050; Upf2; 1.
 DR SMART; SM00543; MIF4G; 2.
 DR Hypothetical protein.
 KW SEQUENCE 596 AA; 69788 MW; F70AD5DD09E165A0 CRC64;

Alignment Scores:
 Pred. No.: 0.118 Length: 596
 Score: 121.00 Matches: 41
 Percent Similarity: 47.89% Conservative: 27
 Best Local Similarity: 28.87% Mismatches: 48
 Query Match: 7.37% Indels: 26
 DB: 2 Gaps: 6

AF334735 (1-954) x Q9GM47 (1-596)

QY 265 GAGACCTCTTAGAGAAAAGAGAGAAACCACTTTGATCCAGCAGATCGGGAGTAG 324
 Db 439 GluAsnLeuGluGluAspGluGlu-----GluGluGlyGlyAlaGlu 453

QY 325 GTAGAGACCCCTCTTATAACAATCATCGATTCGAGGAGCAAGACCACCTGAGAAAGT 384
 Db 454 ThrGluGluGlnSerGlyAsnGluSerGluValAsnGluProGluGluGluGlySer 473

QY 385 GATCTAAACAAGAGAGCTCTCAGATATCTGGGAGGAGGAGAGACATCAGTCACCATC 444
 Db 474 AspAsnAspAspAspGluGlyGlu-----GluGluGluGluGluAsnThrAspTyrLeu 491

QY 445 TTAGACTCTTCTGAGGAAGATAGAAAAGAGAGGTTGCTGCTGTCAAAATCCAAGCT 504
 Db 492 ThrAspSerAsnLysGluAsnGluThrAspGluGluSerThrGlyValMetIleLysGly 511

QY 505 GCCTCCGGGGACATATA-----GCCAGAGAGGAGGCAAGAAATAACAAACAAAT--- 555
 Db 512 GlyGlyLeuLysHisValProCysValGluGluGlnAlaLysArgGlnGlnGluGlu 531

QY 556 ---AGTCTCTAAATAGGAAAAAGAGAAACCAAGTCAGGACACTGGTTTTACCTCCAG 612
 Db 532 AlaAlaAlaGlnMetLysGluLysGluGluSer----- 542

QY 613 GAAACATGAAATAATCAATCCATCACTCTTATTATTAATGTCATTCTTCTCTGAGG 672
 Db 543 -----ValGlnLeuHisGlnGluAlaTrpGluArgHis---HisLeuArg 556

```

QY 673 AACGAA 678
DB 557 LysGlu 558

RESULT 22
Q6BU41
ID Q6BU41 PRELIMINARY; PRT; 562 AA.
AC Q6BU41;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similarities with CA4297|IPF2603 Candida albicans IPF2603.
GN ORFNames=DEHA0C14828g;
OS Debaryomyces hansenii CBS767.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomycetes.
OX NCBI_TaxID=284592;
RN [1]_TaxID=284592;
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RG Genolevures;
RA Lafontaine I., de Montigny J., Marck C., Neugeglise C., Talla E.,
RA Coffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boissrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Ostas S., Ozier-Kalogeropoulos O.,
RA Pellenn S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swenne D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RN Nature 430:35-44(2004).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; C382135; CAG86356.1; -
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; Bromodomain; 1.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 1.
DR PROSITE; PS50014; BROMODOMAIN 2; 1.
SQ SEQUENCE 562 AA; 64900 MW; 395B482BC6EAD44F CRC64;

Alignment Scores:
Pred. No.: 0.159 Length: 562
Score: 119.50 Matches: 40
Percent Similarity: 41.46% Conservative: 28
Best Local Similarity: 24.39% Mismatches: 70
Query Match: 7.28% Indels: 26
DB: Gaps: 4

AF334735 (1-954) x Q6BU41 (1-562)
QY 108 GGTTCATGAGCGAGTTCCTTACCAAGAAAGATGTC-----GATTCCATTCCTCAACAC 158
DB 185 GlyGluValGlnLysAlaTyrGluGlnAspIleGluAlaGluAspGlnGluAsp 204
QY 159 CCATCTACCGAATTCACAGAGATTTGGGAATCTTCTTGAAGGGCTGACACGGGAGATCTT 218
DB 205 LysIleGlyLeuAspValGluValGluHisAspHisAsnGluLysAspIleGluAspIle 224
QY 219 GAGAGAGACACCGGACAAATATACCGCTTTTCGACAGCCTATTTTGAGAGCCTTCTAGA 278
DB 225 Glu-GluHisGluAsnAsn-----GluLeuAs 233
QY 279 GAAAGAGAGAAACCAACCACTTTTGATCCAGAGATGGGGGAGTAGGTAGAGACCGCTT 338

```

```

DB 233 pLysGluGluThrGluAsnIleGluGlnGlyAspAsnIleGluGluGluGlu 253
QY 339 CTATAACAATCATGCAATTCGAGGACGAAGAACCCTCGAGAAAGTAGTCCTTAAACAAGA 398
DB 253 nValGluAsn-----LysAspGluAsnProAspGluProGluAspLysGluGlnGlu 270
QY 399 AGAGTCTCAGATATCTGGGAAGAGGAGAGACATCAGTCACCATCTTTAGACTCTTCTCA 458
DB 270 uGluAspLysGluGlnLysGluGlnGluGluGluGluGluGluGluGluGluGlu 282
QY 459 GGAACATATAGGAAAAAGAGAGAGAGTTCCTCTCAAAATCCAAAGCTGCCTCCGGGGACA 518
DB 282 uGluAspLysGluGlnGluGlnGluGlnGluGlnGluGlnGluGlnGluGlnGluGlu 302
QY 519 CATAGCCAGAGAGAGGAGGCAAGAAATGAAACAAATAGTCTTCAAAATGAGGAAACAA 578
DB 302 nGluGluGlnGluGlnGluGlnGluGlnGluGlnGluGlnGluGlnGluGlnGluGln 322
QY 579 GGAACAAACAG 588
DB 322 uLysGluGlu 325

RESULT 23
Q9BUB1
ID Q9BUB1 PRELIMINARY; PRT; 382 AA.
AC Q9BUB1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE PKAR2A protein (protein kinase, CAMP-dependent, regulatory, type II, alpha).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalilus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA Koundinya M., Raphael J., Moreira D., Kelley T., LeBaer J., Lin Y.,
RA Phelan M., Farmer A.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002763; AA02763.1; -
DR EMBL; BT007225; AAP35889.1; -

```



```

FT BINDING      337      337      CAMP 2.
FT BINDING      346      346      CAMP 2.
SQ SEQUENCE     403 AA; 45387 MW; FC4D17DC1C4EE7D6 CRC64;

Alignment Scores:
Pred. No.:      0.188      Length:      403
Score:          118.50     Matches:    38
Percent Similarity: 41.26%  Conservative: 21
Best Local Similarity: 26.57% Mismatches: 53
Query Match:    7.22%     Indels:    31
DB:             1         Gaps:      4

AF334735 (1-954) x KAP2_HUMAN (1-403)
QY 157 ACCCATCTACGAAATCCACAGGATTGGGAATCTCTTTGAAGGGCTGACACCGAGATT 216
   ||||| :|||:||||| ||| |||||:||||| ||| |||||:|||||
Db 1 SerHisIleGlnIleProGlyLeuThrGluLeuGlnGlyTyrThrValGluVal 20
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 217 CTGAGAGACACCGGACCAATATACAGCTTTTGCAGCAGCCTATTTTGAGACCTCTTA 276
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21 LeuArgGlnGlnProProAspLeuValGluPheAlaValGluTyrPheThrArgLeuArg 40
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 277 GAGAAAGAGAGAAACCACTTGTATCCAGCAGATGGGGAGTAGTAGACCGC 336
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 41 GluAlaArgAlaProAlaSerValLeuProAla-----AlaThrProArg 55
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 337 TTTCTATAACAATCATGCTTCAGGAGCAAGAACCACTTGAGAAAGTGTCTTAACAA 396
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 56 GlnSerLeuGlyHisProProGluProGlyProAspArgValAlaAspAlaGly 75
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 397 GAAGAGTCTCAGATATCTCGGAAGGAGGAAGACATCATGTCACCATCTTTCATCT 456
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 76 Asp-----SerGluSer 79
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 457 GAGGAGATAGAGAAAGAGAGGTGTCTGTCTCAAAATCAAGCTCCCTCCGGGA 516
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 80 GluGluAspGluAspLeuGlu-----ValProValProSerArgPheAsnArg 95
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 517 CACATAGCCAGAGAGGAGGCAAGAAATGAAACAAATAGTCTTCAAAATGAGGAAAA 576
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 96 ArgValSer-----ValCysAlaGluThrTyrAsnProAspGluGlu 109
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 577 GAGGAAAAAC 585
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 110 GluGluAsp 112
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 25
QY5160 PRELIMINARY; PRT; 216 AA.
ID Q7S160;
AC Q7S160;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE Predicted protein.
GN Names=NCU09105.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OK74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Iankiev P., Pedersen D., Nelson M., Washburne M.,
RA Seltrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamysysselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Kryzsofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmari S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,

```

```

RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Laufer E.S., Nussbaum C., Birren B.,
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000462; EAA29078.1; -
SQ SEQUENCE 216 AA; 24311 MW; 48000280CBC923CC CRC64;

Alignment Scores:
Pred. No.:      0.266      Length:      216
Score:          116.50     Matches:    38
Percent Similarity: 45.65%  Conservative: 25
Best Local Similarity: 27.54% Mismatches: 37
Query Match:    7.10%     Indels:    38
DB:             2         Gaps:      4

AF334735 (1-954) x Q7S160 (1-216)
QY 202 CTGACACCGGAGATTCTGAGAGACCAACCGGACATATATACCAGCTTTTGCACGCGCTAT 261
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8 LeuProLeuAspIleIleLeuAspLysAsnAspSer----- 19
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 262 TTTGAGAGCTTCTAGAGAAAGAGAGAAACCACTTTTGATCCAGCAGAAATGGGGAGT 321
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 20 LysGluAspAsnLysGluAspAsnLysGluGluAspLysGluGluAspLys 39
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 322 AAGTAGAGACCGCTTCTATATAAATCATGCTTCGAGGAGCAAGAACCTGAGAAA 381
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 40 GluGluGluAspLys-----GluGluGluAspLysGluGlu 51
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 382 AGTGATCTTAACACAGAGAGTCTCAGATATCTGCGAGGAGGAGAGACATCAGTCACC 441
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 52 GluAsnLysGluGluGluAspLysGluGluAspLysGluGlu----- 67
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 442 ATCTTAGAGCTTCTCTGAGGAAAGATAGGAAAGAGAGGTTGTCTCTCAAAATCCAA 501
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 68 -----AspLysGluGluAspLysGluGluAsn----- 78
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 502 GCTGCTTCCGGGGACACATGCCAGAGAGAGGAGGCAAGAAATGAAACAATAGTCTT 561
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 79 -----LysGluGluAspLysGluGluGluAsnLysGlu 90
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 562 CAAATAGGAGAAAGAGAAACAGTGAGGACACTGTTTACTCCAGG 613
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 91 GluAspLysGluGluGluGluAsnLys-GluAspThrAlaAlaThrProLys 107
   |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 26
QY5160 PRELIMINARY; PRT; 401 AA.
ID Q8K1M2;
AC Q8K1M2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CAMP-dependent protein kinase type II-alpha regulatory chain.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Testis;
RA Ziembra A.J., Collard M.W.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DDJ databases.
DR EMBL; AF533978; AAM97689.1; -
DR HSP; P12367; IRL2.
DR GO; GO:0005952; C:cAMP-dependent protein kinase complex; IEA.
DR GO; GO:0008603; F:cAMP-dependent protein kinase regulator act. . .; IEA.
DR GO; GO:0006301; F:kinase activity; IEA.
DR GO; GO:0006469; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.

```

DR InterPro; IPR002373; CAMP kin.
 DR InterPro; IPR000595; CNMP_binding.
 DR InterPro; IPR003117; RIIa.
 DR Pfam; PF00027; CNMP_binding; 2.
 DR Pfam; PF02197; RIIa; 1.
 DR PRINTS; PR00103; CAMPKINASE.
 DR SMART; SM00100; CNMP; 2.
 DR SMART; SM00394; RIIa; 1.
 DR PROSITE; PS00888; CNMP_BINDING_1; 2.
 DR PROSITE; PS00889; CNMP_BINDING_2; 2.
 DR PROSITE; PS00442; CNMP_BINDING_3; 2.
 DR Kinase.
 SQ SEQUENCE 401 AA; 45540 MW; 026FEBFCFACA9C7D CRC64;

Alignment Scores:
 Pred. No.: 0.28 Length: 401
 Score: 116.50 Matches: 38
 Percent Similarity: 42.66% Conservative: 23
 Best Local Similarity: 26.57% Mismatches: 49
 Query Match: 7.10% Indels: 33
 DB: 2 Gaps: 5

AF334735 (1-954) x Q8KIM2 (1-401)

QY 157 ACCCACTACCGAATTCACAGGATTTGGGAATCTTCTTGAAGGCTGACACGCGAGATT 216
 Db 2 SerHisIleGlnIleProProGlyLeuThrGluLeuGlnGlyThrValGluVal 21
 QY 217 CTGAGAGCAACCGGACAAATATACAGCTTTTGAGCAGCCCTATTGAGAGCCCTCTTA 276
 Db 22 LeuArgGlnProProAspLeuValAspPheAlaValGluThrPheThrArgLeuArg 41
 QY 277 GAGAAAGAGAGAAACCAACTTTGATCCAGCAGATGGGGAGTAAGGTAGAACCGC 336
 Db 42 GluAlaArgArgGlnGluSer 50
 QY 337 TTTCTATAACAATCATGCAATTCGAGGACGAAGAACCACTGAGAAAGTGCATCTAAACAA 396
 Db 51 PheIle 61
 QY 397 GAAGAGTCTCAGATATCTGGGAGGAGGAGGAGCATGTCACCATCTTAGACTCTTCT 456
 Db 62 GluSerSerGlyValProValIleGluGluAsp 75
 QY 457 GAGGAAGATAAGGAAAGAGAGAGTGTGCTGTCAAAATCCAAAGCTCCCTCCGGGA 516
 Db 76 GluSerAspSerAspGluAspLeu 94
 QY 517 CACATAGCCAGAGAGGAGGACAAAGAAATGAAACAAATAGTCTTCAAAATGAGGAAAAA 576
 Db 95 ArgValSer 108
 QY 577 GAGGAAAC 585
 Db 109 GluAspAsn 111

RESULT 27

ID Q710D7 PRELIMINARY; PRT; 498 AA.
 AC Q710D7;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Sperm protein 13.
 OS Vulpes vulpes (Red fox).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Vulpes.
 OX NCBI_TaxID=9627;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Verdier Y.;
 RT "Selection, identification et caracterisation partielle d'antigenes

RT spermatisques de renards.";
 RL Thesis (2002), Department of Biological Sciences, Universite Nancy 1,
 RL Nancy, France.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Verdier Y., Farre G., Rouet N., Boue F.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ421969; CAD19163.1;
 DR GO; GO:0008603; P:camp-dependent protein kinase regulator act. .; IEA.
 DR GO; GO:0007165; F:signal transduction; IEA.
 DR InterPro; IPR003117; RIIa.
 DR Pfam; PF02197; RIIa; 1.
 DR SMART; SM00394; RIIa; 1.
 SQ SEQUENCE 498 AA; 53087 MW; CFFD605645421BA32 CRC64;

Alignment Scores:
 Pred. No.: 0.285 Length: 498
 Score: 116.50 Matches: 34
 Percent Similarity: 48.62% Conservative: 19
 Best Local Similarity: 31.19% Mismatches: 37
 Query Match: 7.10% Indels: 19
 DB: 2 Gaps: 4

AF334735 (1-954) x Q710D7 (1-498)

QY 169 ATTCACAAGGATTTGGGAATCTTCTTGAAGGCTGACACGCGAGATTCTGAGAGCA 228
 Db 10 ValProTyrGlyLeuLeuThrGluLeuGlnGlyValSerArgAlaIleLeuLysIleAsn 29
 QY 229 CCGAGCAATATACCAAGCTTTTGACAGCAGCCCTATTGAGAGCCCTTCTAGAGAAAGAG 288
 Db 30 ProProAsnIleThrGlnPheAlaAlaValTyrPheLysGluLeuIleValPheArgGlu 49
 QY 289 ---AAACCAACTTGTATCCAGCAGAAATGGGGAGTAGGTAGAGACCCGCTTCTATAC 345
 Db 50 GlyAsnThrSerLeuAsp 62
 QY 346 AATCATGCTATTCGAG 378
 Db 63 PheHisGlnIleLysValGluLysTrpSerGluGlyThrThrGlnGluLysGluProGlu 82
 QY 379 AAAAGTGATCTCTCAACAGAGAGTCTCAGATATCTGGGAGGAGGAGAGACATCAGTC 438
 Db 83 CysMetGluGluGlnValGluThrSerValValSer 100
 QY 439 ACCATCTTAGACTCTCTGAGAGAGAT 465
 Db 101 GluLysSerThrAspThrGluGluAsp 109

RESULT 28

KAP2_BOVIN
 ID KAP2_BOVIN STANDARD; PRT; 400 AA.
 AC P00515;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE CAMP-dependent protein kinase type II-alpha regulatory subunit.
 GN Name=PRKAR2A;
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Heart muscle;
 RX MEDLINE=8222175; PubMed=6283532;
 RA Takio K., Smith S.B., Krebs E.G., Walsh K.A., Titani K.;
 RT "Primary structure of the regulatory subunit of type II CAMP-dependent
 protein kinase from bovine cardiac muscle.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:2544-2548(1982).
 RN [2]

3D-STRUCTURE MODELING.
 MEDLINE-87157645; PubMed-3030405;
 Weber I.T., Steitz T.A., Bubis J., Taylor S.S.;
 "Predicted structures of CAMP binding domains of type I and II
 regulatory subunits of CAMP-dependent protein kinase.";
 Biochemistry 26:343-351(1987).
 CC -I- SUBUNIT: The inactive form of the enzyme is composed of two
 regulatory chains and two catalytic chains. Activation by CAMP
 produces two active catalytic monomers and a regulatory dimer that
 binds four CAMP molecules.
 CC -I- TISSUE SPECIFICITY: Four types of regulatory chains are found: I-
 alpha, I-beta, II-alpha, and II-beta. Their expression varies
 among tissues and is in some cases constitutive and in others
 inducible.
 CC -I- PTM: A second phosphorylation site has not been located.
 CC -I- SIMILARITY: Belongs to the CAMP-dependent kinase regulatory chain
 family.
 CC -I- SIMILARITY: Contains 2 cyclic nucleotide-binding domains.
 DR FIR; A06618; ORK02R.
 DR PDB; 2APK; Model; @=1-400.
 DR InterPro; IPR002373; CAMP_kin.
 DR InterPro; IPR000595; CNMP_binding.
 DR InterPro; IPR003117; RIIa.
 DR Pfam; PF00027; CNMP_binding; 2.
 DR Pfam; PF02197; RIIa; 1.
 DR PRINTS; PR00103; CAMPKINASE.
 DR SMART; SM00100; CNMP; 2.
 DR SMART; SM00394; RIIa; 1.
 DR PROSITE; PS00888; CNMP_BINDING_1; 2.
 DR PROSITE; PS00889; CNMP_BINDING_2; 2.
 DR PROSITE; PS00890; CNMP_BINDING_3; 2.
 DR 3D-structure; Acetylation; Direct protein sequencing;
 KW Dimerization; Acetylation; Repeat.
 KW Multigene family; Phosphorylation; Repeat.
 FT DOMAIN 1 134 Dimerization and phosphorylation.
 FT NP_BIND 135 256 CAMP 1.
 FT NP_BIND 257 400 CAMP 2.
 FT MOD_RES 1 1 N-acetylserine.
 FT MOD_RES 95 95 Phosphoserine (by PKA).
 FT BINDING 204 204 CAMP 1.
 FT BINDING 213 213 CAMP 1.
 FT BINDING 334 334 CAMP 2.
 FT BINDING 343 343 CAMP 2.
 FT HELIX 144 149
 FT TURN 150 150
 FT STRAND 154 156
 FT TURN 158 159
 FT STRAND 163 163
 FT TURN 165 166
 FT STRAND 168 168
 FT STRAND 171 172
 FT STRAND 174 174
 FT STRAND 177 177
 FT STRAND 179 184
 FT STRAND 192 198
 FT STRAND 202 202
 FT TURN 205 206
 FT STRAND 212 212
 FT STRAND 216 217
 FT STRAND 223 225
 FT STRAND 228 229
 FT STRAND 230 236
 FT HELIX 237 239
 FT TURN 241 241
 FT HELIX 242 263
 FT HELIX 266 270
 FT TURN 271 272
 FT STRAND 277 278
 FT TURN 280 281
 FT STRAND 283 285
 FT STRAND 290 291
 FT STRAND 293 298
 FT STRAND 304 306
 FT STRAND 311 311

FT STRAND 316 316
 FT STRAND 321 322
 FT TURN 328 329
 FT TURN 332 333
 FT HELIX 334 336
 FT STRAND 341 342
 FT STRAND 346 348
 FT STRAND 353 359
 FT HELIX 360 367
 FT TURN 368 369
 FT TURN 371 371
 FT HELIX 372 378
 FT TURN 379 379
 FT HELIX 380 390
 SQ SEQUENCE 400 AA; 44962 MW; 47809DD900B244E7 CRC64;
 Alignment Scores:
 Pred. No.: 0.341 Length: 400
 Score: 115.50 Matches: 46
 Percent Similarity: 39.58% Conservative: 30
 Best Local Similarity: 23.96% Mismatches: 61
 Query Match: 7.04% Indels: 55
 DB: 1 Gaps: 8
 AF334735 (1-954) x KAP2_BOVIN (1-400)
 QY 157 ACCCACTACCGAATTCACAGGATTGGGAATCTCTTGAAGGCTGACGCGAGATT 216
 Db 1 SerHisIleGlnIleProGlyLeuThrGluLeuLeuGlnGlyTyThrValGluVal 20
 QY 217 CTGAGAGAGCAACCGACAATATACAGCTTTTGAGCAGCGCTATTTGAGAGCTTTCTA 276
 Db 21 LeuArgGlnArgProAspLeuValAspPheIaValAspTyThrArgLeuArg 40
 QY 277 GAGAAAGAGAGAGAAACC-----AACCTTT 300
 Db 41 GluAlaArgSerArgAlaSerThrProAlaIaProProSerGlySerGlnAspPhe 60
 QY 301 GATCCA-----GCAGAAATGGGGAGTAAGCTAGAGAC----- 333
 Db 61 AspProGlyAlaGlyLeuValAlaAspAlaValAlaAspSerGluSerGluAspGluGlu 80
 QY 334 -----CGCTTC----- 339
 Db 81 AspLeuAspValProIleProGlyArgPheAspArgValSerValCysAlaGluThr 100
 QY 340 TATACAAATCATGCTTCGAGGAGCAAGAACCCAGCTGAGAAAGTGATCCTAAACAGAA 399
 Db 101 TyrAsnProAspGluGluAspThrAspProArgValIleHisProIleThrAsp 120
 QY 400 GAGTCTCAGATATCTGGGAAGGAGGAGACATCAGTCACCATCTTAGACTCTTCTGAG 459
 Db 121 GlnGlnArgCysArgLeuGlnGluAlaCysLysAspIleLeuLeuPhe----- 136
 QY 460 GAAGATAAGGAAAAAGAGAGTCTGCTGTCAAAATCCAAAGCTTCCTCCGGGACAC 519
 Db 137 LysAsnLeuAspProGluGlnLeuSerGlnVal-----LeuAspAlaMetPhe----- 152
 QY 520 ATAGCCAGAGAGGAGCAAGAAATAGTCTTCAAAATGATGCTTCAAAATGAGGAGAAAGAG 579
 Db 153 -----GluArgThrValLysValAspGluHisValIleAspGlnGlyAspAspGly 169
 QY 580 GAAAAAAGTGGAGGACACTGGTTTACCTCCAGGAA 615
 Db 170 AspAsn-----PheTyrValIleGlu 176
 RESULT 29
 Q9T069
 ID Q9T069 PRELIMINARY; PRT; 532 AA.
 AC Q9T069;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

```

DE Hypothetical protein AT4g37820 (Hypothetical protein
DE AT4g37820/T28119.100).
GN Name=AT4g37820; Synonyms=At4g37820/T28119.100;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Robben J., Grymonprez B., Volckaert G., Bancroft I.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Robben J., Grymonprez B., Volckaert G., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA Hayashizaki Y., Shinzaki K.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AL035709; CAB38930.1; -
DR EMBL; AL161592; CAB80447.1; -
DR EMBL; AK117758; BAC42406.1; -
DR PIR; T06029; T06029.
KW Hypothetical protein.
SQ SEQUENCE 532 AA; 59135 MW; 0E2B76572C5D0DE0 CRC64;

Alignment Scores:
Pred. No.: 0.385 Length: 532
Score: 115.00 Matches: 50
Percent Similarity: 41.3% Conservatives: 29
Best Local Similarity: 26.1% Mismatches: 29
Query Match: 7.01% Indels: 22
DB: 2 Gaps: 6

AF334735 (1-954) x Q9T069 (1-532)
QY 70 ACAACGGAAACCGGCGACAGCTCGAGAGAAAGAGGTTCCATAGCAGTTCTTACC 129
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
226 ThrThrGluAsnGlySerAspGlyGluGlnGlnGluThrLysSerGluLeuAspSer 245
QY 130 AAGAAGATGTCGATTCCTCCATCTCCACACCCACTACCGAATTCACAA-----GGA 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
246 LysThrGlyGlyGlyGlyGlyPheSerAspSerAsnGlyGluLeuProGluThrAsnLeuSer 265
QY 181 TTTTGGAAFTCTTCTTGAAGGCTGACACCGAGATCTTCGAGAGACCAACCGACAATATA 240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
266 ThrSerAsnAlaThrGlu-----ThrThrGluSerSerGlySerAspGluSerGlySer 283
QY 241 CCAGCTTTTGACGACGCTTATTTTGAGCCCTTCTAGAGCCTTCTAGAGAAAGAGAGAAACCACTTT 300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
284 SerGlyLysSerThrGlyThrGlnGlnThrLysAsnGluGluAspGluLysGluLysVal 303
QY 301 GATCCAGCAAGATGGGGAGTAGGTAGAGACCGCTTCTATAAC----- 345
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
304 GlnSerSerGluGluGluSerLysValLysGluSerGlyLysAsnGluLysAspAlaSer 323
QY 346 ----AATCATGCAATTCAGAGCAAGAACCACTCTGAGAAAGTAGCTCTAAACAAGAGAG 402
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
324 SerSerGlnAspGluSerLysGluGluLysProGluAtgGlyLysGluLysSerSer 343
QY 403 TCTCATGATCTGGGAGAGAGAA-----GAGCATCATCTACC 441
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
344 SerGlnGlyGluGlyLysGluGluProGluLysArgGluLysGluAspSerSerSer 363

```

```

QY 442 ATCTAGACTCTTCTGAGGAAGATAAGAAAAAGAGAGTTGTCTGCTGCTCAAAATCAA 501
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
364 GlnGluGluSerLysGluGluProGluAsnLysGluLysGluLysGluLysSerSerGln 383
QY 502 GCTGCTTCCGGGACACATGACGAGAGAGGAGGCAAGAAATAAATAAATAAGTAGTTT 561
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
384 Glu-----GluAsnGluLeuLysGluThrGluLysGluLysGluLysGluLysSerSer 401
QY 562 CAA-----AATGAGAAAGAGGAAACAAAG 588
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
402 GlnGluGlyAsnGluAsnLysGluThrGluLys 412

RESULT 30
Q75SR9 PRELIMINARY; PRT; 992 AA.
ID Q75SR9;
AC Q75SR9;
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Predicted protein.
GN Name=NCU05774.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kanal M., Kamysseilis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Krystofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Piamann M., Sella S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Eboole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000353; EAA30798.1; -.
SQ SEQUENCE 992 AA; 111858 MW; 7443B6FDBDCDD182 CRC64;

Alignment Scores:
Pred. No.: 0.493 Length: 992
Score: 114.00 Matches: 37
Percent Similarity: 45.99% Conservatives: 26
Best Local Similarity: 27.01% Mismatches: 38
Query Match: 6.95% Indels: 36
DB: 2 Gaps: 6

AF334735 (1-954) x Q75SR9 (1-992)
QY 220 AGAGAGCAACCGGCAATATACCACTTTTGCAGCAGCCTATTTTGAGAGCCTCTAGAG 279
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
382 ArgGluAspAlaAspMetIleProGlu-----Ser 391
QY 280 AAAAGAGAGAAACCAACTTTGATCCAGCAAGATGGGGAGTAGAGTAGACCGCTTC 339
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
392 GluArgGluGluGluValAlaAspAlaGlyGln-----ThrAspValGluLysArgAsp 409
QY 340 TAT-----AACAAATCATGTCATTCGAG 360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
410 TyrGluGluArgAspGluMetLysSerSerGlySerSerSerSerSerSerGlyGluAsp 429
QY 361 GAGCAAGAACCACTCTGAGAAAAAGTAGTCCTCTAAACAAGAGAGTCTCAGATATCTGGGAG 420

```


RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 [4]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 [5]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 [6]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 [7]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 [8]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 [9]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 [10]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 [11]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Wilson R.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 [12]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RG WormBase Consortium;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 5 WD repeats.
 DR EMBL; AC006733; AF60491.2; -
 DR WormBase; WBGene00021316; Y32H12A.8.
 DR WormPep; Y32H12A.8; CE28516.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 5.
 DR SMART; SM00320; WD40; 6.
 DR PROSITE; PS0294; WD REPEATS REGION; 1.
 KW Hypothetical protein; Repeat; WD repeat.
 SQ SEQUENCE 3901 AA; 454378 MW; 7759EC84F925D0D CRC64;

Alignment Scores:
 Pred. NO.: 0.74 Length: 3901
 Score: 112.50 Matches: 38
 Percent Similarity: 43.88% Conservative: 23
 Best Local Similarity: 27.34% Mismatches: 37
 Query Match: 6.86% Indels: 41
 DB: 2 Gaps: 7

AF334735 (1-954) x Q9NE53 (1-3901)

QY 208 CCGAGATTCGAGAGCAACCGGACATATACCGCTTTTGGCAGCAGCCTATTGAG 267
 Db 979 GlnAspValIleThrGluGlnHisAspAenIleLysGlyThrSer-----Glu 994
 QY 268 AGCCTCTAGAGAAAGAGAGAAACCAACTTTGTCCAGCAGAGATGGGGGAGTAAGGTA 327

Db 995 AspGluValGluAspArgGluAsp----- 1002
 QY 328 GAAGACCGC-----TTCTATAACAATCATGCTTCGAGGAGCAAGAACCA 372
 Db 1003 GluAspArgThrAspLeuValLysTyrTyrLeuAsn-----ProProPro 1017
 QY 373 CTGTGAGAAAGTGCATCTTAAACAAGACAGAGTCTCAGATATCTCTGGG----- 417
 Db 1018 ProProLysAsnGluValArgValGlnGluGlnGluGlnGluGlySerArgLysArgLys 1037
 QY 418 -----AAGGAGGAGAGACATCATCATCATCTTTAGACATCTTCTGAGGAAGAT 465
 Db 1038 LeuGluGlnIleGlnGluAspGluAspGluGluLeuLeuAspGluGluGluGlu 1057
 QY 466 AAGGAGGAGAGAGAGTTGCTGCTGCTCAAAATCCAGCTGCTCCGGGGGACATAGCC 525
 Db 1058 PheGluGluGluGlu-----GluValAspGluGluTyrGluGlyGluVal--- 1072
 QY 526 AGAGAGGAGGCAAGAAATGAAACAAATAGTCTTCAAAATGAGGAGGAGGAGGAA 582
 Db 1073 ---GluGluValHisArgGluSerSerGluGluLeuAspGluGluGluGluGlu 1090
 RESULT 33
 Q9NE71
 ID Q9NE71 PRELIMINARY; PRT; 845 AA.
 AC Q9NE71;
 DT 01-OCT-2002 (TRENBLrel. 22, Created)
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE ABCF1 protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Small D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the ABC transporter family.
 DR EMBL; BC034488; AAH34488.1; -
 DR HSSP; P58301; 1118.
 DR Genew; HGNC:70; ABCF1.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
 DR GO; GO:0000166; F:nucleotide binding; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR003593; AAA_ATPase.

QY 316 GGGAGTAAAGTAAAGACCGCTTCTATAC-----AATCATGCAATC----- 357
 DB 208 ThrThrGluValLysSerGluArgGluAlaGluValSerHisThrTyrlleleLys 227
 QY 358 -----GAGGACGACAGACCACTGTGAGAAAGTATGCTTAAACAGAGAG 402
 DB 228 AlaThrThrGlyGlyLysLysLysHisGluGluGluLysLysLysLysLys 247
 QY 403 TCTCAGATATCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 459
 DB 248 GluThrLysSerLysLysLysLysLysLysLysLysLysLysLysLysLys 267
 QY 460 GAAGATAAGGAAAAAGAGAGGTTGCTGTCTGCTCAAAATCAAGCTGCTTCCGGGACAC 519
 DB 268 GluAspAspGluSerSerGluHisGlyAlaIleValLeuArgLysAlaPheSerArg 287
 QY 520 ATAGCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 570
 DB 288 AsnGlyAlaValArgThrLysLysGlyLysLysLysLysLysLysLysLys 304
 RESULT 37
 Q8K1M3
 ID Q8K1M3 PRELIMINARY; PRT; 402 AA.
 AC Q8K1M3;
 DT 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
 DE CAMP-dependent protein kinase type II-alpha regulatory chain (Protein
 DE kinase, cAMP dependent regulatory, type II alpha).
 GN Name=Prkar2a;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Testis;
 RA Mowen K.A., Collard M.W.;
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Brownstein M.J., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green B.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Director MGC Project;

RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF533977; AAM97688.1; -
 DR EMBL; BC075623; AAH75623.1; -
 DR EMBL; BC080276; AAH80276.1; -
 DR HSSP; P12367; 1R2A.
 DR MGD; MGI:108025; Prkar2a.
 DR GO; GO:0005515; F:protein binding; IPI.
 DR InterPro; IPR002373; CAMP_kin.
 DR InterPro; IPR000595; CNMP_binding.
 DR InterPro; IPR003117; Riia_binding.
 DR InterPro; IPR010916; TONB_Box_N.
 DR Pfam; PF00027; CNMP_binding; 2.
 DR Pfam; PF02197; Riia; 1.
 DR PRINTS; PR00103; CAMPKINASE.
 DR SMART; SM00100; CNMP; 2.
 DR SMART; SM00394; Riia; 1.
 DR PROSITE; PS00888; CNMP_BINDING_1; 2.
 DR PROSITE; PS00889; CNMP_BINDING_2; 2.
 DR PROSITE; PS00042; CNMP_BINDING_3; 2.
 DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
 KW Kinase.
 SQ SEQUENCE 402 AA; 45589 MW; 505FDD8B35F1BA00 CRC64;
 Alignment Scores:
 Pred. No.: 0.831 Length: 402
 Score: 111.00 Matches: 45
 Percent Similarity: 38.86% Conservative: 30
 Best Local Similarity: 23.32% Mismatches: 62
 Query Match: 6.76% Indels: 56
 DB: 2 Gaps: 7
 AF334735 (1-954) x Q8K1M3 (1-402)
 QY 157 ACCACTACCGAATTCACAGGATTTGGGAATCTTTGAGGGCTGACACGCGAGATT 216
 DB 2 SerHisIleGlnIleProAlaGlyLeuThrGluLeuLeuGlnGlyTyThrValGluVal 21
 QY 217 CTGAGAGAGCAACCGGACATATACAGCTTTTGACGAGCGCTATTTGAGAGCTTCTTA 276
 DB 22 LeuArgGlnGlnProProAspLeuValAspPheAlaValGluTyThrPheThrArgLeuArg 41
 QY 277 GAGAAAAGAGAGAAA-----ACCACTTT----- 300
 DB 42 GluAlaArgArgGlnGlnSerAspThrPheIleValSerProThrThrPheHisThrGln 61
 QY 301 -----GATCCAGCAGATGGGGAGTAAAGTAGAGAC 333
 DB 62 GluSerSerAlaValProValIleGluAspGlyGluSerAspSerAspSerGluAsp 81
 QY 334 -----CGC 336
 DB 82 AlaAspLeuGluValProValProSerLysPheThrArgArgValSerValCysAlaGlu 101
 QY 337 TCTTATACAAATCATGATTCAGAGGAGCAAGACCACTGTGAGAAAGTATCTTAAACAA 396
 DB 102 ThrPheAsnProAspGluGluGluAspAspProArgValValHisProLysThr 121
 QY 397 GAAGAGTCTCAGATATCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 456
 DB 122 AspGluGlnArgCysArgLeuGlnGlnAlaCysLysAspIleLeuLeuPhe----- 138
 QY 457 GAGGAAGATAAGGAAAAAGAGGTTGCTGTCTGCTCAAAATCAAGCTGCTTCCGGGGA 516
 DB 139 ---LysAsnLeuAspGlnGlnLeuSerGlnVal---LeuAspAlaMetPheGlu--- 155
 QY 517 CACATAGCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 576
 DB 156 -----LysIleValLysThrAspGluHisValIleAspGlnGly 168
 QY 577 GAGGAAAACAAAGTAGGACACTGGTGTTCCTCCAGGAA 615
 DB 169 AspAsp-----GlyAspAsnPheTyThrValIleGlu 178

RESULT 38

Q9CYV3 PRELIMINARY; PRT; 134 AA.

AC Q9CYV3; 17, Created

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 26, Last annotation update)

DE Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length

DE enriched library, clone:2810440D03 product:10, 11 DAYS EMBRYO CDNA,

DE RIKEN FULL-LENGTH ENRICHED LIBRARY, CLONE:2810440D03, FULL INSERT

DE SEQUENCE, full insert sequence.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

[1]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;

CA Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.";

RT Meth. Enzymol. 303:19-44(1999).

[2]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;

CA RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection.";

RT Nature 409:685-690(2001).

[3]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

CA The FANTOM Consortium;

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RT Nature 420:563-573(2002).

[4]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;

CA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,

CA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to

RT prepare full-length cDNA libraries for rapid discovery of new genes.";

RT Genome Res. 10:1617-1630(2000).

[5]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;

CA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,

CA Konno H., Akiyama J., Nishi K., Kikunai T., Tashiro H., Itoh M.,

CA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,

CA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,

CA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,

CA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunura S., Kawai J.,

CA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "RIKEN integrated sequence analysis (RISA) system-384-format

RT sequencing pipeline with 384 multicapillary sequencer.";

RT Genome Res. 10:1757-1771(2000).

[6]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

CA Adachi J., Aizawa K., Akhira S., Akimura T., Arai A., Aono H.,

CA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,

CA Hanagaki T., Hara A., Hayatsu M., Hiramoto K., Hiraoka T., Hori F.,

CA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,

CA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,

CA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,

CA Okazaki Y., Okido T., Owa C., Saio H., Saio R., Sakai C., Sakai K.,

CA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,

CA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,

CA Teijina Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,

CA

Muramatsu M., Hayashizaki Y.;

Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

EMBL; AK013274; BAB28762.1; -.

DR HSSP; P12367; IR2A.

DR GO; GO:0008503; F:CAMP-dependent protein kinase regulator act. . . ; IEA.

DR GO; GO:0007165; P:signal transduction; IEA.

DR InterPro; IPR003117; RIIA.

DR InterPro; IPR010916; TONB_Box_N.

DR Pfam; PF02197; RIIA; 1.

DR SMART; SM00394; RIIA; 1.

DR PROSITE; PS00430; TONB_DEPENDENT_RBC_1; UNKNOWN 1.

SQ SEQUENCE 134 AA; 15229 MW; 19C837606B6548A8 CRC64;

Alignment Scores:

Pred. No.: 0.841 Length: 134

Score: 110.50 Matches: 31

Percent Similarity: 42.74% Conservative: 19

Best Local Similarity: 26.50% Mismatches: 52

Query Match: 6.73% Indels: 15

DB: 2 Gaps: 2

AP334735 (1-954) x Q9CYV3 (1-134)

Qy 157 ACCCACTACCAAGATTCCCAAGATTGGGAATCTTCTGAAGGGCTGCACCGCAGATT 216

Db 2 SerHisileGlnileProAlaGlyLeuThrGluLeuGlnGlyThrValGluVal 21

Qy 217 CTGAGAGACGACCGGACAATATACCAAGCTTTTGGCAGCAGCCCTATTTTGGAGCCCTCTA 276

Db 22 LeuArgGlnProProAspLeuValAspPheAlaValGluThrPheThrArgLeuArg 41

Qy 277 GAGAAAGACGAGAAACCAACTTTGATCCACGAGAAATGGGGAGTAGGTAGAGAGACCGC 336

Db 42 GluAlaArgArgGlnGluSer-----AspThr 50

Qy 337 TTTCTATACAAAT-----CATGCATTTCAGAGACGAGAACACCTGTGAGAAAGT 394

Db 51 PheIleValSerProThrPheHisThrGlnGluSerSerAlaValProValIleGlu 70

Qy 385 GATCTCTAAACAGNAGAGTCTCAGATATCTGGGAGGAGGAGACATCAGTCACCATC 444

Db 71 GluAspGlyGluSerAspSerAspSerGluAspAlaAspLeuGluAspProIleAlaCys 90

Qy 445 TTAGACTCTCTGAGGAGGAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 495

Db 91 LeuLeuSerGluAspLysSerThrSerThrArgGlnLeuAlaProAsnLys 107

RESULT 39

Q8QTD2 PRELIMINARY; PRT; 921 AA.

AC Q8QTD2;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE WSSV315.

OS White spot syndrome virus (WSSV).

OC Viruses; dsDNA viruses, no RNA stage; Nimaviridae; Whispovirus.

OX NCBI_TaxID=92652;

RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE=20517548; PubMed=11062040; DOI=10.1006/viro.2000.0597;

RA Tsai M.F., Yu H.T., Tzeng H.F., Leu J.H., Chou C.M., Huang C.J.,

RA Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;

RT "Identification and characterization of a shrimp white spot syndrome

RT virus (WSSV) gene that encodes a novel chimeric polypeptide of

RT cellular-type thymidine kinase and thymidylate kinase.";

RL Virology 277:100-110(2000).

[2]

RP SEQUENCE FROM N.A.

RC MEDLINE=21844071; PubMed=11853398; DOI=10.1006/viro.2001.1273;

RA Chen L.L., Leu J.H., Huang C.J., Chou C.M., Chen S.W., Wang C.H.,

RA Lo C.F., Kou G.H.;

RT "Identification of a nucleocapsid protein (VP35) gene of shrimp white

RT

RT spot syndrome virus and characterization of the motif important for
RT targeting VP35 to the nuclei of transfected insect cells.";
RL Virology 293:44-53 (2002).

[3]
RN SEQUENCE FROM N.A.
RP Lo C.-P., Kou G.-H.;
RA Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RL
DR EMBL; AF440570; AAL89183.1; --
SQ SEQUENCE 921 AA; 102933 MW; C483AE887175C345 CRC64;

118	GCAGTCTTCCACGAGAGATCGATTCCATTCCTCCAAACCCACCTACCGAATCCACAA	177
Qy		
121	AlaValAlaGluArgIleAlaMetLysGluAspThrHisIyr	136
Db		
178	CGATTGGGAATCTTCTTGAAGGCTGACACGCGAGATCTTGAGAGAGCAACCGGCAAT	237
Qy		
137	---ValGlyValAlaProLysValLysArgGluAlaThrArgGluArg---	152
Db		
238	ATACAGCGCTTTTGCAGCGCTATTTTTGAGAGCGCTTCTAGAGAAA	282
Qy		
153	---LysAlaLeuAlaAlaLysValGluLysAlaIleGluLysGluLeuLeuAspArg	171
Db		
283	-----AGAGAGAAAACCAACTTTTGATCCAGCAGAGATCGGGGAGT	321
Qy		
172	LeuLysSerGlyAlaTyGlyGluAsnProLeuIleuValAspGluLysValTrpLysLys	191
Db		
322	AAGGTGAAGACCGCTTCTATAACAATCATGCATTCGAGGACGAAGAACCACTTGAAA	381
Qy		
192	ValLeuGluGluValGluGluValAsnLysGluAspGluAspGluGluGluGluGlu	211
Db		

DR	SWART; SM00394; RIIa; 1.
DR	PROSITE; PS00888; CNMP_BINDING_1; 1.
DR	PROSITE; PS00889; CNMP_BINDING_2; 1.
DR	PROSITE; PS00442; CNMP_BINDING_3; 2.
KW	CAMP-binding; Phosphorylation; Repeat.
DNA	155 Dimerization and phosphorylation
FT	DOMAIN 1 155 (Potential).
FT	DOMAIN 90 102 Asp/Glu-rich (acidic).
FT	NP_BIND 156 278 CAMP 1.
FT	NP_BIND 279 403 CAMP 2.
FT	MOD_RES 117 117 Phosphoserine (By similarity).
FT	BINDING 226 226 CAMP 1 (By similarity).
FT	BINDING 235 235 CAMP 1 (By similarity).
FT	BINDING 349 349 CAMP 2 (By similarity).
FT	BINDING 358 358 CAMP 2 (By similarity).
FT	VARIANT 372 372 K -> T.
SQ	SEQUENCE 403 AA; 44467 MW; 4PFFEP48B27B2A40 CRC64;

Alignment Scores:

Pred. No.:	1-5	Length:	403
Score:	108.00	Matches:	32
Percent Similarity:	40.62%	Conservative:	20
Best Local Similarity:	25.00%	Mismatches:	68
Query Match:	6.58%	Indels:	8
DB:	1	Gaps:	1

```

AF334735 (1-954) x KAPR_BLAEM (1-403)
QY      163 TACCGAATTCACAGGATTGGGAATCTTCTTGAAAGGTGCACACGCGAGATCTTGAGA 222
         ||| ||||| ||||| :~::~: ||||| ||||| ||||| |||||
Db       4 TyrThrIleProSerGluLeuProProIleLeuLysAspLeuSerArgGluValLeuArg 23
QY      223 GAGCAACCGGACAATATACCAGCTTTCCAGCAGCCTATTTCAGAGCCCTTTCAGAGCAA 282
         ||| :~::~: ||||| ||||| ||||| ||||| ||||| |||||
Db      24 HisGlnProAlaAspLeuValGlnPheCysHisAspTyrPheAlaLysLeuLeuAlaGln 43
QY      283 AGAGAGAAAACCAACTTTTGATCCAGCAAGATGGGGAGTAGGTAGAGAACGCCGCTTCAT 342
         :~::~: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      44 GinArgLysValLeuMetAspSerAlaAspProAlaThrLysAlaThrIleAlaSerThr 63
QY      343 AACATCATGCATTCAGAGGACGACGACCACTCAGAAAAGTGNCTCTAAACAAGAGAG 402
         ||| :~::~: ||||| ||||| ||||| ||||| ||||| |||||
Db      64 AlaGlyProAlaValAspAlaAspGluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 63
QY      403 TCTCAGATATCTGGGAAGGAGGAGACATCAGTCACCATCTTAGACTCTTCTGAGGAA 462
         ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      84 AspAspGlyPheGlyThrGlu-----AspAspAspAspAsp 95
QY      463 GATAAGGAAAAGAGAGGTTGCTGCTGTCAAAATCCAAGTCGCTTCGGGGACACATA 522
         ||| :~::~: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      96 AspAspGluAspGluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 515
QY      523 GCCAGAGGAGGCAAGAAATG 546
         ||| :~::~: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     116 ThrSerValSerAlaGluSerMet 123

RESULT 45
YKX5_YEAST
ID YKX5_YEAST STANDARD; PRT; 1132 AA.
AC F34250;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypothetical 125.6 kDa protein in AAT1-GFAL intergenic region.
GN OrderedLocusNames=YKU105C; ORFNames=YKU459;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RL [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=94152173; PubMed=8109175;

```



```

Alignment Scores:
Pred. No.: 1.68 Length: 475
Score: 107.50 Matches: 34
Percent Similarity: 44.17% Conservative: 19
Best Local Similarity: 28.33% Mismatches: 36
Query Match: 6.55% Indels: 31
DB: 2 Gaps: 4

AF334735 (1-954) x Q8WVW5 (1-475)
QY 169 ATTCCAAAGATTGGGAATCTTTGAAGGCTGACACGCGAGATTCTGAGAGCA 228
D 10 ValProTyrGlyLeuLysThrLeuGluGluGlySerArgAlaValLeuLysThrAsn 29
QY 229 CCGGACNATATACCGCTTTTCAGCAGCCTATTTTGAGGCTTCTAGAGAAAGAGAG 288
D 30 ProSerAsnIleAsnGlnPheAlaAlaAlaTyrPheGlnGluLeu 44
QY 289 AAAACCAACTTTGATCCAGCAAGTGGGGAGTAAAGTAGAGACCGCTTCTATAACAAT 348
D 45 ---ThrMetTyrArgValGluLysTyrSer 53
QY 349 CATGCATTTCAGGAGCAAGAACCCCTGAGAAAGTGATCTCTAAACAAGAA 399
D 54 -----GluGlyThrThrProGlnLysLysLeuGluCysLeuLysGluProGlyLys 70
QY 400 -----GAGTCTCAGATA-----TCTGGGAAGAGGAGAGAGACA 432
D 71 ThrSerValGluSerLysValProThrGlnMetGluLysSerThrAspThrAspGluAsp 90
QY 433 TCAGTCACCATCTTACGATCTTCTCAGGAAGATAGGAAAGAGAGAGGTTCTGCTGTC 492
D 91 AsnValThrArgThrGluTyrSerAspLysThrThrGlnPheProSerValTyrAlaVal 110

RESULT 48
IF2_AQUAE
ID IF2_AQUAE STANDARD; PRT; 805 AA.
AC O67825;
DT 15-DEC-1998 (Rel. 37, Created)
DT 25-DEC-1998 (Rel. 37, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Translation initiation factor IF-2.
GN Name=infB; OrderedLocusNames=AQ_2032;
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320; DOI=10.1038/32831;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aufay M., Huber R.,
RA Feldman R.A., Short J.M., Olsen G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus".
RL Nature 392:353-358(1998).
CC -!- FUNCTION: One of the essential components for the initiation of
CC protein synthesis. Protects formylmethionyl-tRNA from spontaneous
CC hydrolysis and promotes its binding to the 30S ribosomal subunits.
CC Also involved in the hydrolysis of GTP during the formation of the
CC 70S ribosomal complex (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the IF-2 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL; AE000769; AAC07794.1; -.
DR PIR; E70474; E70474.
DR HSSP; O26359; 1G7S.
DR HAMAP; MF_00100; -.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR000178; IF2.
DR InterPro; IPR006847; IF2_N.
DR InterPro; IPR000795; ProSyn_GTPbind.
DR InterPro; IPR005225; Small_GTP.
DR InterPro; IPR009000; translat_factor.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 2.
DR Pfam; PF04760; IF2_N; 2.
DR PRINTS; PR00315; ELONGATNFT.
DR PRODOM; PD18100; IF2; 1.
DR TIGRFAMs; TIGR00487; IF-2; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR PROSITE; PS01176; IF2; 1.
KW Complete proteome; GTP-binding; Initiation factor;
KW protein biosynthesis.
FT DOMAIN 309 458 G-domain.
FT NP_BIND 315 322 GTP (By similarity).
FT NP_BIND 362 366 GTP (By similarity).
FT NP_BIND 416 419 GTP (By similarity).
SQ SEQUENCE 805 AA; 91600 MW; 8A1BF300C69CA51E CRC64;

Alignment Scores:
Pred. No.: 1.76 Length: 805
Score: 107.50 Matches: 53
Percent Similarity: 42.49% Conservative: 29
Best Local Similarity: 27.46% Mismatches: 74
Query Match: 6.55% Indels: 37
DB: 1 Gaps: 8

AF334735 (1-954) x IF2_AQUAE (1-805)
QY 73 ACCGAAACGGCGCCACCGAGCTCGAGAGAAAGAGGAGTTCATAGGAGTTCCTACCAAG 132
D 4 ThrLysThrLysArgValSerValAlaLysGluLeu-----GlyValLysSerLys 21
QY 133 AAGATGTCGATTCATTCCTCCACACCCACTACCGAATTCACCAAGGA-----180
D 22 GluIle---IleGluPheLeuAsnGluTyrTyrProArgProAspGlyLysProTyrLys 40
QY 181 -----TTGGGAATCTT 192
D 41 AlaSerHisGlyLeuAspGluGlnAlaLeuGluMetIleTyrAspAlaPheGlyIleLys 60
QY 193 CTTGAGGGCTGCACACGCGAGATTCTGAGAGAGCAACCGGACATATATACCGAGCTTTTGA 252
D 61 GluGluGluLysGluGluValValThrGluGln-----AlaGlnAla 75
QY 253 GCAGCCTATTTCGAGAGCCTTCTAGAGAGAAAGAGAGAAACCAACTTTGATCCAGCAGAA 312
D 76 ProAlaGluValGlu-----GluLysLysGluGluLysLysGluGluValLe 92
QY 313 TGGGGAGTAGTAGTAGAAGACCGCTTCTATACATCATGCATTCGAGAGGAGCAAGACCA 372
D 93 ValGluGluValValGluGluLys---LysProGluValIleValGluGluLysGlu 111
QY 373 CCTGAGAGAAAGTAGTCTCTAACAGAGAGCTCTCAGATATCTCGGAGAGGAGGAGACACA 432
D 112 LysLysGluGluGluLysLysLysGluGluLysProLysLysSerValGluGluLeu 131
QY 433 TCAGTCACCATCTTACGACTCTTCTGAGAGAAAGAGAGAAAGAGAGTTCGCTGCTC 492
D 132 IleLysGluIleLeuGluLysLysGluLysGluLysGluLysLysValGluLysGlu 151
QY 493 AATATCAAGCTGCTTCCTCCG---GGACACATGCCAGAGAGAGGAGGCAAGAAATGAAA 549
D 152 ArgLysGluLysValArgValValGluValLysLysGluLysGluLysGluLys 171
QY 550 ACAAATAGTCTTCAAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 588

```

```

Db 172 -----LysGluGluLysLysGluGluGluLys 180
RESULT 49
Q7RK82
ID Q7RK82 PRELIMINARY; PRT; 881 AA.
AC Q7RK82;
DT 01-WAR-2004 (TREMELrel. 26, Created)
DT 01-WAR-2004 (TREMELrel. 26, Last sequence update)
DT 01-WAR-2004 (TREMELrel. 26, Last annotation update)
DE Maeb1 (Fragment).
GN Name:PY03020;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OC NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=1236865; DOI=10.1038/nature01099;
RA Carlton J.M., Anguoli S.V., Suh B.B., Kooij T.W., Perteu M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallow S.J., van Aken S.B., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoabi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RA "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii."
RT Nature 419:512-519 (2002).
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; ABL01000854; EAA22546.1; -.
DR HSSP; P17676; IHJB.
FT NON TER 881
SQ SEQUENCE 881 AA; 105238 MW; F6B81ADD18C98B79 CRC64;

Alignment Scores:
Pred. No.: 1.77 Length: 881
Score: 107.50 Matches: 50
Percent Similarity: 46.88% Conservative: 40
Best Local Similarity: 26.04% Mismatches: 82
Query Match: 6.55% Indels: 20
DB: 2 Gaps: 9

AF334735 (1-954) x Q7RK82 (1-881)
QY 39 AAGAAAACGGTTACCCAGCACTAGAAAAACAAACCGGCGGCGCAGCAGCTCGGA 98
Db 285 LysLysLysGluGluAspValAspPheAlaAsnLeuArgAsnLysLysGluAsp 304
QY 99 GAGAAAGAGGTTCATAGCGGATCTTACCAAGAGATGTCGATTCATTCACACAC 158
Db 305 ValAspPheAlaAsnLeuArgAsnLysLysGluGlu-AspValAspPheSerAsnLe 324
QY 159 CCNACTACCAATTCACAA-----CGATTGGGAATCTCTTGAAGGCGTGACA-- 207
Db 324 uArgAsnLysLysLysGluGluAspValAspPheGlyAsnLeuArgAsnLysLysAsnAs 344
QY 208 -CGCAGATCTCGAGAGACCAACCGGCAATATACCGCTTTTGCAGCAGCCTATTTTCA 266
Db 344 pThrGluLysLysLysLysLysLysProTyrHisLeuArgArgLeuGluGluLeuArgIleG 364
QY 267 GAGCCTCTTAGAGAAAGAGAGAAACCAACTTTGATCCAGAGATGGGGGAGTAAGGT 326
Db 364 nLysGlnLeu---LysArgGlnGluThrGluLysSerGln-----AsnGlyLysLys 380
QY 327 AGAAGACCGCTCTATACAAATCATCGATTCGAGGAGCAAGAACCACTGAG-----AA 380
Db 380 uGluGluLysLysLysLysLysLysGluAspGluLysLysLysLysValGluGluLysLe 400

RESULT 50
Q8I590
ID Q8I590 PRELIMINARY; PRT; 3193 AA.
AC Q8I590;
DT 01-WAR-2003 (TREMELrel. 23, Created)
DT 01-WAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-WAR-2003 (TREMELrel. 23, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PFL1600C;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OC NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2255705; PubMed=1236864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Hall A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.S., Nene V., Shallow S.J., Suh B., Peterson J., Anguoli S.,
RA Perteu M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RA "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RT Nature 419:498-511 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Hyman R.W., Fung E., Conway A., Kurdi O., Mao J., Miranda M.,
RA Nakao B., Rowley D., Tamaki T., Wang F., Davis R.W.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DDJ databases.
DR EMBL; AB014849; AAN36406.1; -.
KW Hypothetical protein.
SQ SEQUENCE 3193 AA; 377925 MW; 46E6B21F921C5307 CRC64;

Alignment Scores:
Pred. No.: 1.96 Length: 3193
Score: 107.50 Matches: 51
Percent Similarity: 39.52% Conservative: 32
Best Local Similarity: 24.29% Mismatches: 63
Query Match: 6.55% Indels: 64
DB: 2 Gaps: 10

AF334735 (1-954) x Q8I590 (1-3193)
QY 39 AAGAAAACGGTTACCCAGCACTAGAAAAACAAACCGGCGGCGCAGCAGCTCGGA 98
Db 1641 LysLysAsnAspValThrAspAsnAsnGluAsnAsnGlyLysArgAsnLysLysLeuArg 1660
QY 99 -----GAGAAAGAGGTTCACATA-----GCG 119
Db 1661 LysAlaLysThrPheGlyGlyThrValLysLysGlyGlyThrIleLeuAsnLeuGluAla 1680
QY 120 AGTTCTTACCAAGAGATGTCGATTCATTCCTCC-----AACAC 158

```

Db 1681 GluSer-GlnArgIleLysArgThrProPheCysIleTyrSerLeuAsnAspIleCysAs 1700
QY 159 CCACCTACCGAATTCACAGGATTTGGGAATCTTCTGAGGGCTGACACGC----- 210
Db 1700 nHisLeuArgTyrGlyLysLysSerAlaAsnSerIleAsnGlyMetIleLysThrSerTh 1720
QY 211 -----GAGATTCTGAGAGACCAACCGGACAATATACCAGCTTTTCAGCAGCCTA 260
Db 1720 rCysAsnCysGluIle-----HisAsnAspMetLysPheValLysGlnTy 1735
QY 261 TTTTGAGAGCCTTCAGAGAAAGAGAGAAACCAACTTTGATCCAGCAGAAATGGGGAG 320
Db 1735 rIleAsnPheLeuIleAlaGlnLysGluLysGluArgIleAsn----- 1749
QY 321 TAAGGTAGAACCCCTTCTATAACAATCATGCTTCGAGGAGCAAGAACCACTGAGAA 380
Db 1750 -----LysGluLysLysIleAsn-----GluLysGluLysAsnAspG1 1763
QY 381 AAGTCATCCTAAACAAGAGTCTCAGATATCTGGGAAGGAGGAGAGACATCAGTCAC 440
Db 1763 uGlyAsnLysLysGluGluGluSerGln-----LysGluGluGluSer----- 1777
QY 441 CATCTTAGACTTCTGAGGAGATPAGGAAAGAGAGGTTGCTGCTGTCAAAATCCA 500
Db 1778 -----GlnLysGluGluGluSerGlnLysGluGluSerLys----- 1790
QY 501 AGCTGCCTTCGGGGACACATAGCCAGAGAGAGGCAAGAAATGAACAAATAGTCT 560
Db 1791 -----AsnGluGluGluSerLysAsnGluGluAsnAsnIleLy 1803
QY 561 TCAAAATCAGGAAAGAGGAAACAAAG 588
Db 1803 sAspAsnGluGlnLysArgLysAsnLys 1812

Search completed: February 15, 2005, 10:06:29
Job time : 189.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 15, 2005, 02:05:31 ; Search time 201 Seconds
(without alignments)
4860.931 Million cell updates/sec

Title: AF334735

Perfect score: 307

Sequence: 1 TCGCCCTTCGCGCGCG.....AAAAAAAAAAAAAAAAAAAA 954

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3224408

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q/cgn2_1/USPTO.spool_p/AF334735/runat_14022005_160237_19339/app_query.fasta_1.1095
-DB=UniProt_03 -Qfmt=fastan -SUFFIX=oligo.rup -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human4g/cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=AF334735 @CGN 1 1 244 @runat_14022005_160237_19339 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : UniProt 03.*

1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	151	49.2	151	1 SP17 HUMAN	Q15506 homo sapien
2	52	30.0	151	1 SP17 MACFA	Q19021 macaca fasc
3	91	29.6	163	1 SP17 PAPHA	Q95230 papio hamad
4	71	23.1	143	2 Q9BDQ6	Q9BDQ6 macaca mula
5	44	14.3	146	1 SP17 RABIT	P36425 oryctolagus
6	44	14.3	148	2 Q9Z1K2	Q9Z1K2 rattus norv
7	44	14.3	149	1 SP17 MOUSE	Q62252 mus musculu
8	44	14.3	153	1 SP17 MACEU	Q62770 macropus eu
9	44	14.3	179	1 SP17 MONDO	Q62771 monodelphis
10	40	13.0	147	2 Q9TU08	Q9TU08 ovis aries
11	33	10.7	212	2 Q9XT29	Q9XT29 callithrix
12	15	4.9	141	2 Q6SJ92	Q6SJ92 ornithorhyn
13	13	4.2	495	2 Q6DIJ9	Q6DIJ9 xenopus tro
14	11	3.6	1012	2 Q7S3H9	Q7S3H9 neurospora
15	10	3.3	1148	2 Q8FN75	Q8FN75 corynebacte
16	10	3.3	1448	2 Q9XZE9	Q9XZE9 hydra atten

17	9	2.9	139	2	Q7QHMI	Q7qhml anopheles g
18	9	2.9	145	1	STIA_XENLA	Q09006 xenopus lae
19	9	2.9	471	2	Q73MIS	Q73mi5 treponema d
20	9	2.9	549	2	Q9G910	Q9g910 ochromonas
21	9	2.9	863	2	Q9BTP1	Q9bcp1 homo sapien
22	9	2.9	1040	2	Q8NHN2	Q8nnh2 homo sapien
23	9	2.9	6620	2	Q96AA2	Q96aa2 homo sapien
24	8	2.6	48	2	Q7M2U5	Q7m2u5 bubalus bub
25	8	2.6	48	2	Q8F422	Q8f422 leptospira
26	8	2.6	66	2	Q49241	Q49241 mycoplasma
27	8	2.6	67	2	Q72W40	Q72w40 leptospira
28	8	2.6	67	2	Q8F9S8	Q8f9s8 leptospira
29	8	2.6	88	2	Q61WB2	Q61wb2 arabidopsis
30	8	2.6	102	2	Q8C5Y8	Q8c5y8 mus musculu
31	8	2.6	105	2	Q35129	Q35129 neurospora
32	8	2.6	110	2	Q7QP28	Q7qp28 giardia lam
33	8	2.6	112	2	Q8BYC1	Q8byc1 mus musculu
34	8	2.6	112	2	Q9CUV5	Q9cuv5 mus musculu
35	8	2.6	123	2	Q653R8	Q653r8 oryza sativ
36	8	2.6	133	2	Q6PIM9	Q6pim9 homo sapien
37	8	2.6	139	1	YED3_YEAST	P32633 saccharomyc
38	8	2.6	140	2	Q8BRM4	Q8brm4 mus musculu
39	8	2.6	143	2	Q9C4A7	Q9c4a7 tricholoma
40	8	2.6	150	2	Q8ENS8	Q8ens8 oceanobacil
41	8	2.6	151	2	Q9AQY4	Q9aqy4 guillardia
42	8	2.6	158	2	Q88B26	Q88b26 pseudomonas
43	8	2.6	166	2	Q9BGP4	Q9bgp4 macaca fasc
44	8	2.6	170	2	Q9NXR4	Q9nxr4 homo sapien
45	8	2.6	177	2	Q8L3Q3	Q8l3q3 oryza sativ

ALIGNMENTS

RESULT 1

SP17_HUMAN
ID SP17_HUMAN STANDARD; PRT; 151 AA.
AC Q15506; Q9BXF7;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Sperm surface protein Sp17 (Sperm autoantigenic protein 17) (Sperm protein 17) (Sp17-1).
GN Name=SPAL7; Synonyms=SP17;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=96305346; PubMed=8688458; DOI=10.1016/0167-4781(96)00077-2;
RA Lea I.A., Richardson R.T., Widgren E.E., O'Rand M.G.;
RT "Cloning and sequencing of cDNAs encoding the human sperm protein, Sp17.";
RL Biochim. Biophys. Acta 1307:263-266(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22280349; PubMed=12393185; DOI=10.1016/S0167-4781(02)00478-5;
RA Buchli R., De Jong A., Robbins D.L.;
RT "Genomic organization of an intron-containing sperm protein 17 gene (Sp17-1) and an intronless pseudogene (Sp17-2) in humans: a new model.";
RL Biochim. Biophys. Acta 1578:29-42(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler N.K., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore S.I., Wang J., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

Alignment Scores:
 Pred. No.: 1.23e-85 Length: 151
 Score: 92.00 Matches: 92
 Percent Similarity: 100.00% Conservations: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 29.97% Indels: 0
 DB: 1 Gaps: 0

AF334735 (1-954) x SP17_PAPHA (1-151)

QY 136 ATGTCGATTCATCTCCACACCCACTACCGAATTCACAGATTGGGAATCTCTT 195
 DB 1 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeu 20
 QY 196 GAAGGGCTGACCGCAGATTCTGAGAGCAACCGACATATACAGCTTTTGCAGCA 255
 DB 21 GluGlyLeuThrArgGluLeuLeuArgGluGlnProAsnIleProAlaPheAla 40
 QY 256 GCCTATTTTTCAGAGCCCTTCTAGAGAAAGAGAGAAACCAATCTTTCATCCAGCAGATGG 315
 DB 41 AlaTyrPheGluSerLeuLeuGluLysArgGluLysThrAsnPheAspProAlaGluTTP 60
 QY 316 GCGAGTAAGTAGAAGCCCTCTATACATCATGCTTCGAGAGCAAGAACCACTT 375
 DB 61 GlySerLysValGluAspArgPheTyrAsnAsnHisAlaPheGluGlnGluProPro 80
 QY 376 GAGAAAGTGCATCTTAAACAGAGAGAGTCTCAGATA 411
 DB 81 GluLysSerAspProLysGlnGluGluSerGlnLe 92

RESULT 3
 SP17_PAPHA STANDARD; PRT; 163 AA.
 AC Q95230;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Sperm surface protein Sp17 (Sperm autoantigenic protein 17).
 GN Names=SP17; Synonyms=SP17;
 OS Papio hamadryas (Hamadryas baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Papio.
 OX NCBI_TaxID=9557;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=97264490; PubMed=9110316;
 RA DOI=10.1002/(SICI)1098-2795(199705)47:1<66::AID-MRD9>3.0.CO;2-O;
 RA Adoyo P.A., Lea I.A., Richardson R.T., Widgren E.E., O'Rand M.G.;
 RT "Sequence and characterization of the sperm protein Sp17 from the
 RT baboon.";
 RL Mol. Reprod. Dev. 47:66-71(1997).
 CC -!- FUNCTION: Sperm surface zona pellucida binding protein. Helps to
 CC bind spermatozoa to the zona pellucida with high affinity. Might
 CC function in binding zona pellucida and carbohydrates (By
 CC similarity).
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Membrane-associated (Potential).
 CC -!- TISSUE SPECIFICITY: Testis- and sperm-specific.
 CC -!- SIMILARITY: Contains 1 IQ domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U75209; AAB38534.1; -;
 CC HSSP; P12367; IR2A.

DR InterPro; IPR000048; IQ region.
 DR InterPro; IPR003117; RIIA.
 DR Pfam; PF0612; IQ; 1.
 DR SMART; PF02197; RIIA; 1.
 DR SMART; SM00015; IQ; 1.
 DR SMART; SM00394; RIIA; 1.
 DR PROSITE; PS00096; IQ; 1.
 KW Membrane.
 FT DOMAIN 114 143 IQ.
 SQ SEQUENCE 163 AA; 18755 MW; A890B59B3F6469BD CRC64;

Alignment Scores:
 Pred. No.: 1.33e-84 Length: 163
 Score: 91.00 Matches: 91
 Percent Similarity: 100.00% Conservations: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 29.64% Indels: 0
 DB: 1 Gaps: 0

AF334735 (1-954) x SP17_PAPHA (1-163)

QY 136 ATGTCGATTCATCTCCACACCCACTACCGAATTCACAGATTGGGAATCTCTT 195
 DB 1 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeu 20
 QY 196 GAAGGGCTGACCGCAGATTCTGAGAGCAACCGACATATATACAGCTTTTGCAGCA 255
 DB 21 GluGlyLeuThrArgGluLeuLeuArgGluGlnProAsnIleProAlaPheAla 40
 QY 256 GCCTATTTTTCAGAGCCCTTCTAGAGAAAGAGAGAAACCACTTTCATCCAGCAGATGG 315
 DB 41 AlaTyrPheGluSerLeuLeuGluLysArgGluLysThrAsnPheAspProAlaGluTTP 60
 QY 316 GCGAGTAAGTAGAAGCCCTCTATACATCATGCTTCGAGAGCAAGAACCACTT 375
 DB 61 GlySerLysValGluAspArgPheTyrAsnAsnHisAlaPheGluGlnGluProPro 80
 QY 376 GAGAAAGTGCATCTTAAACAGAGAGAGTCTCAG 408
 DB 81 GluLysSerAspProLysGlnGluGluSerGln 91

RESULT 4
 Q9BDQ6 PRELIMINARY; PRT; 143 AA.
 AC Q9BDQ6;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Sperm protein 17 (Fragment).
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=22280349; PubMed=12393185; DOI=10.1016/S0167-4781(02)00478-5;
 RA Buchli R., De Jong A., Robbins D.L.;
 RT "Genomic organization of an intron-containing sperm protein 17 gene
 RT (Sp17-1) and an intronless pseudogene (Sp17-2) in humans: a new
 RT model.";
 RL Biochim. Biophys. Acta 1578:29-42(2002).
 DR EMBL; AF334809; AAK28124.1; -;
 DR HSSP; P12367; IR2A.
 DR GO; GO:0008603; FcAMP-dependent protein kinase regulator act. . . ; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR000048; IQ region.
 DR InterPro; IPR003117; RIIA.
 DR Pfam; PF0612; IQ; 1.
 DR Pfam; PF02197; RIIA; 1.
 DR SMART; SM00015; IQ; 1.
 DR SMART; SM00394; RIIA; 1.

```
DR PROSITE; PS50096; IQ; 1.
FT NON_TER 1
FT NON_TER 143
SQ SEQUENCE 143 AA; 16424 MW; 70B538F7F876B465 CRC64;

Alignment Scores:
Pred. No.: 7,48e-64 Length: 143
Score: 71.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 23.13% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x Q9BDQ6 (1-143)
QY 154 AACACCACCTACCGAATTCACAGGATTGGGAATCTCTTGAAGGGCTGACACCGGAG 213
Db 1 AenThrHisTyrArgIleProGlnGlyPheGlyAsnLeuGluGlyLeuThrArgGlu 20
QY 214 ATTCGAGAGACCAACCGACAAATATACAGCTTTTGCAGCAGCCTATTTCGAGAGCCTT 273
Db 21 IleLeuArgGluGlnProAspAsnIleProAlaPheAlaAlaAlaTyrPheGluSerLeu 40
QY 274 CTAGAGAAAGAGAGAAAACCACTTTGATCCAGCAGAAATGGGGAGTAGAGTAGAAGAC 333
Db 41 LeuGluLysArgGluLysThrAsnPheAspProAlaGluTrpGlySerLysValGluAsp 60
QY 334 CGCTTCCTAATCAATCATGCTTCGAGAGACAA 366
Db 61 ArgPheTyrAsnAsnHisAlaPheGluGluGln 71

RESULT 5
SPI7_RABIT
ID SPI7_RABIT STANDARD; PRT; 146 AA.
AC F36425;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Sperm surface protein SPI7 (Sperm autoantigenic protein 17).
GN Names=SPA17; Synonyms=SPI7;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 113-126.
RC STRAIN=New Zealand white; TISSUE=Testis;
RX MEDLINE=95046885; PubMed=7525387; DOI=10.1006/dbio.1994.1285;
RA Richardson R.T., Yamaaki N., O'Rand M.G.;
RT "Sequence of a rabbit sperm zona pellucida binding protein and
RT localization during the acrosome reaction.";
RL Dev. Biol. 165:688-701(1994).
CC -!- FUNCTION: Sperm surface zona pellucida binding protein. Helps to
CC bind spermatozoa to the zona pellucida with high affinity. Might
CC function in binding zona pellucida and carbohydrates.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Membrane-associated (Potential).
CC -!- TISSUE SPECIFICITY: testis- and sperm-specific.
CC -!- PFM: The N-terminus is blocked.
CC -!- SIMILARITY: Contains 1 IQ domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 220655; CAA79674.1; -.
CC FIR; 146506; 146506.
CC HSSP; F12367; 1R2A.
CC InterPro; IPR000048; IQ_region.

DR InterPro; IPR003117; RIIA.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF02197; RIIA; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00394; RIIA; 1.
DR PROSITE; PS50096; IQ; 1.
KW Direct protein sequencing; Membrane.
FT DOMAIN 110 139
SQ SEQUENCE 146 AA; 16891 MW; 3FF302B883D9566F CRC64;

Alignment Scores:
Pred. No.: 7,48e-36 Length: 146
Score: 44.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.33% Indels: 0
DB: 1 Gaps: 0

AF334735 (1-954) x SPI7_RABIT (1-146)
QY 136 ATGTGATTCCATTTCTCCACACCCACTACCGAATTCACAGGATTGGGAATCTCTT 195
Db 1 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeu 20
QY 196 GAAGGGCTGACACGGCAGATTCTGAGAGAGCAACCGGACAAATATACCGCTTTTCGAGCA 255
Db 21 GluGlyLeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAlaAla 40
QY 256 GCCTATTTTGAG 267
Db 41 AlaTyrPheGlu 44

RESULT 6
Q9Z1K2
ID Q9Z1K2 PRELIMINARY; PRT; 148 AA.
AC Q9Z1K2;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE SPI7 protein.
GN Name=spi7;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Frayne J., Jury J.A., Hall L.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ131888; CAA10524.1; -.
DR HSSP; F12367; 1R2A.
DR GO; GO:0008603; P:camp-dependent protein kinase regulator act. . .; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR003117; RIIA.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF02197; RIIA; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00394; RIIA; 1.
DR PROSITE; PS50096; IQ; 1.
SQ SEQUENCE 148 AA; 17036 MW; 5E0BC9D4C86PD585 CRC64;

Alignment Scores:
Pred. No.: 7,47e-36 Length: 148
Score: 44.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.33% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x Q9Z1K2 (1-148)
```

```

QY 136 ATGTGATTCATTCCTCCAAACACCCACTACCGAATTCACCAAGGATTTGGGAATCTTCTT 195
Db 1 MetSerIleProPheSerAenThrHisTyArgIleProGlnGlyPheGlyAenLeuLeu 20
QY 196 GAAGGCTGACACCGGATTCCTGAGAGACACCGACCAATATACACGCTTTTGCAGCA 255
Db 21 GluGlyLeuThrArgGluIleLeuArgGluGlnProAspAsnIleProAlaPheAlaAla 40
QY 256 GCCTATTTTGGAG 267
Db 41 AlaTyR-PheGlu 44

RESULT 7
SP17_MOUSE
ID SP17_MOUSE STANDARD; PRT; 149 AA.
AC Q62252;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE Sperm surface protein Sp17 (Sperm autoantigenic protein 17).
GN Name=Spal7; Synonyms=Sp17;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Testis;
RX MEDLINE=96039129; PubMed=7578682;
RA Kong M., Richardson R.T., Widgren E.B., O'Rand M.G.;
RT "Sequence and localization of the mouse sperm autoantigenic protein,
RT Sp17."
RL Biol. Reprod. 53:579-590(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavani T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Tohiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Sperm surface zona pellucida binding protein. Helps to
CC bind spermatozoa to the zona pellucida with high affinity. Might
CC function in binding zona pellucida and carbohydrates (By
CC similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Membrane-associated (Potential).
CC -!- TISSUE SPECIFICITY: Testis- and sperm-specific.
CC -!- SIMILARITY: Contains 1 IQ domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL; Z46299; CA86455.1; -.
DR EMBL; BC059727; AAHS9727.1; -.
PIR; S49527; S49527.
DR HSRP; P12367; IR2A.
DR MGD; MGI:1333778; Spal7.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR003117; RIIA.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF02197; RIIA; 1.
DR PROSITE; PS50096; IQ; 1.
KW Membrane.
FT DOMAIN 112 141 IQ.
SQ SEQUENCE 149 AA; 17296 MW; C7E05D111D6AF0DC CRC64;

Alignment Scores:
Pred. No.: 7,46e-36 Length: 149
Score: 44.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.33% Indels: 0
DB: 1 Gaps: 0

AF334735 (1-954) x SP17_MOUSE (1-149)
QY 136 ATGTGATTCATTCCTCCAAACACCCACTACCGAATTCACCAAGGATTTGGGAATCTTCTT 195
Db 1 MetSerIleProPheSerAenThrHisTyArgIleProGlnGlyPheGlyAenLeuLeu 20
QY 196 GAAGGCTGACACCGGATTCCTGAGAGACACCGACCAATATACACGCTTTTGCAGCA 255
Db 21 GluGlyLeuThrArgGluIleLeuArgGluGlnProAspAsnIleProAlaPheAlaAla 40
QY 256 GCCTATTTTGGAG 267
Db 41 AlaTyR-PheGlu 44

RESULT 8
SP17_MOUSE
ID SP17_MOUSE STANDARD; PRT; 153 AA.
AC Q62770;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Sperm surface protein Sp17 (Sperm autoantigenic protein 17).
GN Name=Spal7; Synonyms=Sp17;
OS Macropus eugenii (Tamar wallaby).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX NCBI_TaxID=9315;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Wen Y., O'Rand M.G.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Sperm surface zona pellucida binding protein. Helps to
CC bind spermatozoa to the zona pellucida with high affinity. Might
CC function in binding zona pellucida and carbohydrates (By
CC similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Membrane-associated (Potential).
CC -!- TISSUE SPECIFICITY: Testis- and sperm-specific (By similarity).
CC -!- SIMILARITY: Contains 1 IQ domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF054289; AAC08024.1; -.

```

DR HSSP; P12367; 1R2A.
 DR InterPro; IPR000048; IQ region.
 DR InterPro; IPR003117; RIIA.
 DR Pfam; PF00612; IQ; 1.
 DR SMART; SM00015; IQ; 1.
 DR SMART; SM00197; RIIA; 1.
 DR SMART; SM00015; IQ; 1.
 DR SMART; SM00394; RIIA; 1.
 DR PROSITE; PS50096; IQ; 1.
 KW Membrane.
 FT DOMAIN 122 151 IQ.
 SQ SEQUENCE 153 AA; 17327 MW; AD3A7CGBD4D4E1B65 CRC64;
 Alignment Scores:
 Pred. No.: 7 43e-36 Length: 153
 Score: 44.00 Matches: 44
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 14.33% Indels: 0
 DB: 1 Gaps: 0

AF334735 (1-954) x SP17_MACEU (1-153)
 QY 136 ATGTGATTCATTCCTCCACACCCACTACCGAATTCACAGGATTTGGGAATCTTCTT 195
 Db 1 MetSerileProPheSerAsnThrHisTyArgIleProGlnGlyPheGlyAsnLeuLeu 20
 QY 196 GAAGGCTCAGCGGAGATTCGAGAGACGACCAACGAGCAATATACCGATTTTGCAGCA 255
 Db 21 GluGlyLeuThrArgGluLeuArgGluGlnProAspAsnIleProAlaPheAlaAa 40
 QY 256 GCCTATTTTGAG 267
 Db 41 AlaTyPheGlu 44

RESULT 9
 SP17_MONDO
 ID SP17_MONDO STANDARD; PRT; 179 AA.
 AC O6271;
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Sperm surface protein Sp17 (Sperm autoantigenic protein 17).
 GN Names:SP17; Synonyms:SP17;
 OS Monodelphis domestica (Short-tailed grey opossum).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Monodelphis.
 OX NCBI_TaxID=13616;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Wen Y., O'Rand M.G.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Sperm surface zona pellucida binding protein. Helps to
 CC bind spermatozoa to the zona pellucida with high affinity. Might
 CC function in binding zona pellucida and carbohydrates (By
 CC similarity).
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Membrane-associated (Potential).
 CC -!- TISSUE SPECIFICITY: Testis- and sperm-specific (By similarity).
 CC -!- SIMILARITY: Contains 1 IQ domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; AF054290; AAC08025.1; -;
 DR HSSP; P12367; 1R2A.
 DR InterPro; IPR000048; IQ region.
 DR InterPro; IPR003117; RIIA.

DR Pfam; PF00612; IQ; 1.
 DR Pfam; PF02197; RIIA; 1.
 DR SMART; SM00015; IQ; 1.
 DR SMART; SM00394; RIIA; 1.
 DR PROSITE; PS50096; IQ; 1.
 KW Membrane.
 FT DOMAIN 124 132 Poly-Glu.
 FT DOMAIN 143 172 IQ.
 SQ SEQUENCE 179 AA; 20864 MW; CADDCl3CEC66A00E CRC64;
 Alignment Scores:
 Pred. No.: 7 27e-36 Length: 179
 Score: 44.00 Matches: 44
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 14.33% Indels: 0
 DB: 1 Gaps: 0

AF334735 (1-954) x SP17_MONDO (1-179)
 QY 136 ATGTGATTCATTCCTCCACACCCACTACCGAATTCACAGGATTTGGGAATCTTCTT 195
 Db 1 MetSerileProPheSerAsnThrHisTyArgIleProGlnGlyPheGlyAsnLeuLeu 20
 QY 196 GAAGGCTCAGCGGAGATTCGAGAGACGACCAACGAGCAATATACCGATTTTGCAGCA 255
 Db 21 GluGlyLeuThrArgGluLeuArgGluGlnProAspAsnIleProAlaPheAlaAa 40
 QY 256 GCCTATTTTGAG 267
 Db 41 AlaTyPheGlu 44

RESULT 10
 Q9TU08
 ID Q9TU08 PRELIMINARY; PRT; 147 AA.
 AC Q9TU08;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE SP17 protein.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Jejunal Peyer's patch;
 RX MEDLINE=21015023; PubMed=11132159; DOI=10.1007/s002510000253;
 RA Tatlow D., Brownlie R., Babiuk L.A., Griebel P.;
 RT "Differential display analysis of gene expression during the induction
 RT of mucosal immunity.";
 RL Immunogenetics 52:73-80 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Jejunal Peyer's patch;
 RA Tatlow D.D., Brownlie R., Babiuk L.A., Griebel P.J.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF179926; AAD54394.1; -;
 DR HSSP; P12367; 1R2A.
 DR GO; GO:0008603; F:camp-dependent protein kinase regulator act. . ; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR000048; IQ region.
 DR InterPro; IPR003117; RIIA.
 DR Pfam; PF00612; IQ; 1.
 DR Pfam; PF02197; RIIA; 1.
 DR SMART; SM00015; IQ; 1.
 DR SMART; SM00394; RIIA; 1.
 DR PROSITE; PS50096; IQ; 1.
 SQ SEQUENCE 147 AA; 16903 MW; 4F9CB44103C25E66 CRC64;
 Alignment Scores:
 Pred. No.: 1 05e-31 Length: 147

OS Ornithorhynchus anatinus (Duckbill platypus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Monotremata; Ornithorhynchidae; Ornithorhynchus.
NCBI_TaxID=9258;
[1]
RN
RP SEQUENCE FROM N.A.
RC
RX TISSUE=Testis;
RC MEDLINE=21308450; PubMed=11415432; DOI=10.1042/0264-6021.3570025;
RA Wen Y., Richardson R.T., Widgren E.E., O'Rand M.G.;
RT "Characterization of Sp17: a ubiquitous three domain protein that
RT binds heparin";
RL Biochem. J. 357:25-31(2001).
[2]
RN
RP SEQUENCE FROM N.A.
RC
RX TISSUE=Testis;
RA Alekseev O. Jr., Richardson R.T., O'Rand M.G.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR ENBL, AY457927; AR21214.1; .
DR GO; GO:0008603; F:CAMP-dependent protein kinase regulator act. . ; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000048; IQ region.
DR InterPro; IPR003117; RIIa.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF02197; RIIa; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00394; RIIa; 1.
DR SMART; PS50096; IQ; 1.
DR PROSITE; PS50096; IQ; 1.
SQ SEQUENCE 141 AA, 15752 MW, 98E3282E0314D0B0 CRC64;

Alignment Scores:
Pred. No.: 8.95e-06 Length: 141
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.89% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x Q6SJ92 (1-141)
QY 223 GAGCACCGGACAAATACCACTTTTGCAGCAGCCTATTTTGAG 267
DB 30 GAGGInProAspAsnIleProAlaPheAlaAlaIaIaTyPheGlu 44

RESULT 13
Q6DIJ9 PRELIMINARY; PRT; 495 AA.
AC Q6DIJ9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Sp17-prov protein.
GN Name=sp17-prov;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
NCBI_TaxID=8364;
[1]
RN
RP SEQUENCE FROM N.A.
RC
RX TISSUE=Embryo;
RC MEDLINE=223388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zesberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalon D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalilov D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Gerhard D.S.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC075539; AH75539.1; -;
 DR GO; GO:0008603; F:AMP-dependent protein kinase regulator act. . ; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR000048; IQ region.
 DR InterPro; IPR003117; RIIa.
 DR Pfam; PF00612; IQ; 2.
 DR Pfam; PF02197; RIIa; 1.
 DR SMART; SM00015; IQ; 2.
 DR SMART; SM00394; RIIa; 1.
 DR PROSITE; PS0096; IQ; 2.
 SQ SEQUENCE 495 AA; 56219 MW; 4B857410B790912A CRC64;

Alignment Scores:
 Pred. No.: 0.00893 Length: 495
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.23% Indels: 0
 DB: 2 Gaps: 0

AF334735 (1-954) x Q6DIJ9 (1-495)

QY 136 ATGTGATTCCTTCACACCCACTACCGAATTCAC 174
 |||||
 Db 1 MetSerIleProPheSerAsnThrHisTyArgIlePro 13

RESULT 14

Q7S3H9 PRELIMINARY; PRT; 1012 AA.
 AC Q7S3H9; (1-MAR-2004 (TRENBLrel. 26, Created)
 DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN Name=NCU06874.1;
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OR74A;
 RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
 RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
 RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
 RA Qui D., Iankiev P., Pedersen D., Nelson M., Washburne M.,
 RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
 RA Kothke G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
 RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
 RA Kamal M., Kamysseilis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
 RA Krystofova S., Raemissen C., Metzberg R.L., Perkins D.D., Kroken S.,
 RA Cognigni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
 RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
 RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
 RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbels D.J., Freitag M.,
 RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
 RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
 RL Nature 0:0-0(2003).
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.
 DR EMBL; AABX01000409; EAA29985.1; -;
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR000910; HMG 12 box.
 DR Pfam; PF00505; HMG box; 1.
 DR PROSITE; PS50118; HMG_BOX_2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 1012 AA; 110158 MW; 56323711A39D7B01 CRC64;

Alignment Scores:

Pred. No.: 0.0959 Length: 1012
 Score: 11.00 Matches: 11
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.58% Indels: 0
 DB: 2 Gaps: 0

AF334735 (1-954) x Q7S3H9 (1-1012)

QY 920 TCTATCATCCCAAAAAAATAAAAAA 952
 |||||
 Db 581 SerIleThrSerProLysLysLysLysLys 591

RESULT 15

Q8FN75 PRELIMINARY; PRT; 1148 AA.
 AC Q8FN75;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Putative ribonuclease.
 GN OrderedLocustNames=CE2272;
 OS Corynebacterium efficiens.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=152794;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YS-314;
 RX MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;
 RX Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
 RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
 RA Gojobori T.;
 RT "Comparative complete genome sequence analysis of the amino acid
 RT replacements responsible for the thermostability of Corynebacterium
 RT efficiens.";
 RL Genome Res. 13:1572-1579(2003).
 DR EMBL; AP005221; BAC19082.1; -;
 DR GO; GO:0005737; C:cytoplasm; IEA.
 DR GO; GO:0004540; F:ribonuclease activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003743; P:translation initiation factor activity; IEA.
 DR GO; GO:0006396; P:RNA processing; IEA.
 DR GO; GO:0006413; P:translational initiation; IEA.
 DR InterPro; IPR006847; IP2 N.
 DR InterPro; IPR008994; Nucleic acid OB.
 DR InterPro; IPR009061; Putativ_DNA_Bind.
 DR InterPro; IPR004659; RNaseEG.
 DR InterPro; IPR003029; S1.
 DR Pfam; PF04760; IP2_N; 1.
 DR Pfam; PF00575; S1; 1.
 DR SMART; SM00316; S1; 1.
 DR TIGRFAMs; TIGR00757; RNaseEG; 1.
 DR PROSITE; PS50126; S1; 1.
 KW Complete proteome.
 SQ SEQUENCE 1148 AA; 126647 MW; C364493697D4DC37 CRC64;

Alignment Scores:

Pred. No.: 1.03 Length: 1148
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 3.26% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x Q8FN75 (1-1148)

QY 65 GAAACACCGGACCGCGCGGCGCAGCT 94

Db 1124 GlulysGlnProGluProAlaProAla 1133

Search completed: February 15, 2005, 02:28:53
Job time : 207 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 15, 2005, 02:13:41 ; Search time 45 Seconds
(without alignments)
4079.591 Million cell updates/sec

Title: AF334735
Perfect score: 307
Sequence: 1 TGCCCTCTCTGCGCGCGG.....AAAAAAAAAAAAAAAAAAAA 954

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Word size: 1
Total number of hits satisfying chosen parameters: 565918
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-O=/cgn2.1/USPTO/spool_p/af334735/runat_14022005_160238_19349/app_query_fasta_1.1095
-DB=PIR_79 -QFMT=fastan -SUFFIX=oligo.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -DIST=45
-NORMALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=AF334735 @CGN 1.1.63 @runat_14022005_160238_19349 -NCPV=6 -ICPV=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DSLOP=6 -DELEXT=7

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	151	49.2	151	2 I38243	zona pellucida bin
2	44	14.3	146	2 I46506	zona pellucida bin
3	44	14.3	149	2 S49527	zona pellucida bin
4	9	2.9	145	2 I51706	stathmin - African
5	9	2.9	296	2 T20726	hypothetical prote
6	8	2.6	41	2 T07274	hypothetical prote
7	8	2.6	48	2 A26334	beta-casein - wate
8	8	2.6	139	2 S30850	hypothetical prote
9	8	2.6	151	2 C90082	hypothetical prote
10	8	2.6	151	2 B90102	hypothetical prote
11	8	2.6	151	2 F90101	hypothetical prote
12	8	2.6	151	2 C90118	hypothetical prote
13	8	2.6	151	2 A90138	hypothetical prote
14	8	2.6	209	2 A59068	beta-casein varian

C 15	8	2.6	222	2 JCI384	beta-casein precu
C 16	8	2.6	222	2 A32979	beta-casein precu
C 17	8	2.6	224	1 KBBOA2	beta-casein precu
C 18	8	2.6	226	1 KBHU	embryogenesis-rela
C 19	8	2.6	226	2 T09567	beta-casein - pig
C 20	8	2.6	232	2 A48384	homotetic protein T
C 21	8	2.6	232	2 A61045	hypothetical prote
C 22	8	2.6	250	2 S19518	cold acclimation p
C 23	8	2.6	262	2 S43953	hypothetical prote
C 24	8	2.6	296	2 E71118	NADH dehydrogenase
C 25	8	2.6	349	2 T12120	hypothetical prote
C 26	8	2.6	368	2 G85587	probable inner mem
C 27	8	2.6	368	2 AD0599	hypothetical prote
C 28	8	2.6	368	2 F90737	hypothetical prote
C 29	8	2.6	368	2 H64815	hypothetical prote
C 30	8	2.6	393	2 T33715	Na+/H+-exchanging
C 31	8	2.6	422	2 C70474	hypothetical prote
C 32	8	2.6	448	2 T30982	hypothetical prote
C 33	8	2.6	493	2 F96696	protein FlN21.12 [
C 34	8	2.6	517	2 T49310	hypothetical prote
C 35	8	2.6	624	2 T20445	hypothetical prote
C 36	8	2.6	629	2 T45796	hypothetical prote
C 37	8	2.6	661	2 T37753	probable disease r
C 38	8	2.6	869	2 A71400	hypothetical prote
C 39	8	2.6	897	2 T21688	hypothetical prote
C 40	8	2.6	975	2 T48107	Ca2+-transporting
C 41	8	2.6	1091	1 PWB922	Ca2+-transporting
C 42	8	2.6	1091	2 S67852	Ca2+-transporting
C 43	8	2.6	1091	2 S25007	probable secreted
C 44	8	2.6	1192	2 A71623	hypothetical prote
C 45	8	2.6	1234	2 C97606	

ALIGNMENTS

RESULT 1

I38243
zona pellucida binding protein Sp17 [similarity] - human
C:Species: Homo sapiens (man)
C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C:Accession: I38243; S52921
R:Richardson, R.T.; Yamasaki, N.; O'Rand, M.G.
Dev. Biol. 165, 688-701, 1994
A:Title: Sequence of a rabbit sperm zona pellucida binding protein and localization dur
A:Reference number: I38243; MUID:95046885; PMID:7525387
A:Accession: I38243
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-151 <RES>
A:Cross-references: UNIPROT:Q15506; EMBL:Z48570; NID:g695580; PIDN:CAA8459.1; PID:g695
R:Lea, I.A.; Richardson, R.T.; Widgren, E.R.; O'Rand, M.G.
submitted to the EMBL Data Library, March 1995
A:Description: Human Sp17: a sperm-zona binding protein.
A:Reference number: S52921
A:Accession: S52921
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-151 <LEA>
C:Superfamily: sperm surface protein Sp17

Alignment Scores:
Pred. No.: 4.44e-145 Length: 151
Score: 151.00 Matches: 151
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.19% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x I38243 (1-151)

QY 136 ATGTGATTCGATTCCTCCACACCCACTACCGAATTCACAGGATTCGGATCTCTT 195
|||||
Db 1 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeu 20
|||||

QY 196 GAAGGCTGACACGGAGATTCTTGAGAGAGCAACCGACAATATACCAGCTTTTCAGCA 255
Db |||||||LeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAlaAa 40
QY 256 GCTATTTTGGAGCCTCTAGAGAAAGAGAGAAACCAACTTTGATCCAGCAGAAATGG 315
Db |||||||LeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAlaAa 40
QY 41 AlaTyPheGluSerLeuLeuGluLysArgGluLeuThrAsnPheAspProAlaGluTrp 60
QY 316 GGGAGTAGGTAGAGACCGCTTCTATTAACAATCATGATTCGAGAGAGCAAGAACCTT 375
Db |||||||LeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAlaAa 40
QY 61 GlySerLysValGluAspArgPheTyAsnAsnHisAlaPheGluGluGlnGluProPro 80
QY 376 GAGAAAGTGATCTTAACAAGAGAGTCTAGATATCTGGAAGAGAGAGAGAGAGATCA 435
Db |||||||LeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAlaAa 40
QY 81 GluLysSerAspProLysGlnGluSerGlnIleSerGlyGluGluGlnGluThrSer 100
QY 436 GTCCACATCTTGAAGTCTTCTGAGGAAGATAAGGAAAGAGAGAGTCTGCTGTCAA 495
Db |||||||LeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAlaAa 40
QY 101 ValThrIleLeuAspSerSerGluGluAspLysGluLysGluValAlaAlaValLys 120
QY 496 ATCCAGAGTCTTCGGGGACACATAGCCAGAGAGGAGGCAAGAAATGAAACAAAT 555
Db |||||||LeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAlaAa 40
QY 121 IleGlnAlaAlaPheArgGlyHisIleAlaArgGluGluAlaLysMetLysThrAen 140
QY 556 ACTCTTCAAATGAGGAAAGAGGAAACAG 588
Db |||||||LeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAlaAa 40
QY 141 SerLeuGlnAsnGluLysGluGluAsnLys 151
RESULT 2
I46506
zona pellucida binding protein Sp17 [similarity] - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 14-Feb-1997 #sequence revision 14-Feb-1997 #text_change 09-Jul-2004
C:Accession: I46506; S58441; S31764
R:Richardson, R.T.; Yamasaki, N.; O'Rand, M.G.
Dev. Biol. 165, 688-701, 1994
A:Title: Sequence of a rabbit sperm zona pellucida binding protein and localization dur
A:Reference number: I38243; MUID:195046885; PMID:7525387
A:Accession: I46506
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-146 <RIC>
A:Cross-references: UNIPROT:P36425; EMBL:Z20655; NID:g479083; PIDN:CAA79674.1; PID:g4790
A:Accession: S58441
A:Status: preliminary
A:Molecule type: protein
A:Residues: 113-126 <RI3>
R:Richardson, R.T.; Yamasaki, N.; O'Rand, M.G.
submitted to the EMBL Data Library, January 1993
A:Description: Sequence of a mammalian sperm surface zona pellucida binding protein and
A:Reference number: S31764
A:Accession: S31764
A:Molecule type: DNA
A:Residues: 1-122, 'DT', <RI2>
A:Status: preliminary
A:Cross-references: EMBL:Z20655
C:Superfamily: sperm surface protein Sp17

Alignment Scores:
Pred. No.: 6.32e-36 Length: 146
Score: 44.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.33% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x I46506 (1-146)

QY 136 ATGTCGATTCCATTCACACACCCACTACCGAATTCACAGGATTTGGGAATCTTCT 195
Db |||||||LeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAlaAa 40
QY 1 MetSerIleProPheSerAsnThrHisTyArgIleProGlnGlyPheGlyAsnLeuLeu 20
QY 196 GAAGGCTGACACGGAGATTCTTGAGAGAGCAACCGACAATATACCAGCTTTTCAGCA 255

Db |||||||LeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAlaAa 40
QY 256 GCCTATTTTGAG 267
Db |||||||LeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAlaAa 40
QY 41 AlaTyPheGlu 44
RESULT 3
S49527
zona pellucida binding protein Sp17 [similarity] - mouse
C:Species: Mus musculus (house mouse)
C:Date: 20-Feb-1995 #sequence revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S49527
R:Kong, M.; Richardson, R.T.; O'Rand, M.G.
submitted to the EMBL Data Library, October 1994
A:Description: Sequence and localization of the mouse sperm protein, Sp17.
A:Reference number: S49527
A:Accession: S49527
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-149 <KON>
A:Cross-references: UNIPROT:Q62252; EMBL:Z46299; NID:g561527; PIDN:CAA86455.1; PID:g5615
C:Superfamily: sperm surface protein Sp17

Alignment Scores:
Pred. No.: 6.31e-36 Length: 149
Score: 44.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.33% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x S49527 (1-149)

QY 136 ATGTCGATTCCATTCACACACCCACTACCGAATTCACAGGATTTGGGAATCTTCT 195
Db |||||||LeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAlaAa 40
QY 1 MetSerIleProPheSerAsnThrHisTyArgIleProGlnGlyPheGlyAsnLeuLeu 20
QY 196 GAAGGCTGACACGGAGATTCTTGAGAGAGCAACCGACAATATACCAGCTTTTCAGCA 255
Db |||||||LeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAlaAa 40
QY 256 GCCTATTTTGAG 267
Db |||||||LeuThrArgGluLeuLeuArgGluGlnProAspAsnIleProAlaPheAlaAa 40
QY 41 AlaTyPheGlu 44

RESULT 4
I51706
stathmin - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 13-Sep-1996 #sequence revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: I51706
R:Maucuer, A.; Moreau, J.; Mechali, M.; Sobel, A.
J. Biol. Chem. 268, 16420-16429, 1993
A:Title: Stathmin gene family: phylogenetic conservation and developmental regulation in
A:Reference number: A47345; MUID:93346387; PMID:8344928
A:Accession: I51706
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-145 <MAU>
A:Cross-references: UNIPROT:Q09006; EMBL:X71431; NID:g397173; PIDN:CAA50562.1; PID:g3971
C:Superfamily: stathmin

Alignment Scores:
Pred. No.: 3.2 Length: 145
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.93% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x I51706 (1-145)

QY 917 CTCTATCATCCGCCCAAAAAA 943
|||||
Db 35 LeuSerIleThrSerProLysLys 43

RESULT 5

T20726
hypothetical protein F10G8.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T20726
R:Basham, V.
submitted to the EMBL Data Library, September 1996
A:Reference number: Z19315
A:Accession: T20726
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-296 <WIL>
A:Cross-references: UNIPROT:Q93456; EMBL:Z80216; PIDN:GN00019; CESP:F1
A:Experimental source: clone F10G8
C:Genetics:
A:Gene: CESP:F10G8.7
A:Map position: 1
A:Introns: 53/1; 129/3; 153/3; 187/3; 209/2; 249/2

Alignment Scores:

Pred. No.:	2.94	Length:	296
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.92%	Indels:	0
DB:	2	Gaps:	0

AF334735 (1-954) x T20726 (1-296)

QY 482 ACTCTCTTTTCCTATCTCTCA 456
|||||
Db 27 ThrSerSerPheSerLeuSerSer 35

RESULT 6

T07274
hypothetical protein 41c - Chlorella vulgaris chloroplast
C:Species: chloroplast Chlorella vulgaris
C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-Jul-2000
C:Accession: T07274
R:Wakaugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Nakas
Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997
A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga Chlo
A:Reference number: Z15985; MUID:97303241; PMID:9159184
A:Accession: T07274
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-41 <WAK>
A:Cross-references: EMBL:AB001684; NID:g2224352; PIDN:BAA57922.1; PID:g2224438
C:Genetics:
A:Genome: chloroplast
C:Keywords: chloroplast

Alignment Scores:			
Pred. No.:	38.9	Length:	41
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.61%	Indels:	0
DB:	2	Gaps:	0

AF334735 (1-954) x T07274 (1-41)

QY 24 TTTTCTTTTCTTAAGAAAAA 47
|||||
Db 6 PhePhePhePheLeuLysLys 13

RESULT 7

A26334
beta-casein - water buffalo (fragment)
C:Species: Bubalus arnee (water buffalo)
C>Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 09-Jul-2004
C:Accession: A26334
R:Petrilli, P.; Pucci, P.; Morris, H.R.; Addeo, F.
Biochem. Biophys. Res. Commun. 140, 28-37, 1986
A:Title: Assignment of phosphorylation sites in buffalo beta-casein by fast atom bombar
A:Reference number: A26334; MUID:87048757; PMID:3778448
A:Accession: A26334
A:Molecule type: protein
A:Residues: 1-48 <PET>
A:Cross-references: UNIPROT:Q7M2U5
C:Superfamily: beta-casein

Alignment Scores:			
Pred. No.:	38.2	Length:	48
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.60%	Indels:	0
DB:	2	Gaps:	0

AF334735 (1-954) x A26334 (1-48)

QY 470 TCCTTATCTTCCTCAGAGTCT 447
|||||
Db 15 SerLeuSerSerSerGluSer 22

RESULT 8

S30850
hypothetical protein YEL033w - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C>Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 09-Jul-2004
C:Accession: S30850; S50511
R:Mulligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor,
submitted to the EMBL Data Library, February 1993
A:Reference number: S30812
A:Accession: S30850
A:Molecule type: DNA
A:Residues: 1-139 <MUL>
A:Cross-references: UNIPROT:P32633; GB:U18779; EMBL:L10830; NID:g603625; PIDN:AAB65009.1
R:Dietrich, F.S.
submitted to the EMBL Data Library, December 1994
A:Description: The sequence of S. cerevisiae cosmids 8199, 8334, and 9871.
A:Reference number: S50491
A:Accession: S50511
A:Molecule type: DNA
A:Residues: 1-139 <DIE>
A:Cross-references: EMBL:U18779; NID:g603625; PIDN:AAB65009.1; PID:g603646; MIPS:YEL033w
C:Genetics:
A:Cross-references: SGD:S0000759
A:Map position: 5L
C:Superfamily: Saccharomyces hypothetical protein YEL033w
C:Keywords: transmembrane protein

Alignment Scores:			
Pred. No.:	33.7	Length:	139
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.61%	Indels:	0
DB:	2	Gaps:	0

AF334735 (1-954) x S30850 (1-139)

QY 931 CCCCCAAAAA954
|||||
Db 103 ProGlnLysLysLysLysLys 110

RESULT 9

C90082

hypothetical protein orf151 [imported] - Guillardia theta nucleomorph
C:Species: nucleomorph Guillardia theta
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: C90082
R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Rei
Nature 410, 1091-1096, 2001
A:Title: The highly reduced genome of an enslaved algal nucleus.
A:Reference number: A99082; MUID:11323671; PMID:11323671
A:Accession: C90082
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-151 <DOU>
A:Cross-references: UNIPROT:Q9AQY4; GB:AF165818; NID:g13794566; PIDN:AAK39941.1; GSPDB:C
C:Genetics:
A:Gene: orf151
A:Map position: 1
A:Genome: nucleomorph
C:Keywords: nucleomorph

Alignment Scores:
Pred. No.: 33.4 Length: 151
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.60% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x C90082 (1-151)

QY 594 TCCTCACTGTTTCTCTTTTC 571

Db 28 SerSerLeuValPheLeuPhePhe 35

RESULT 10

B90102
hypothetical protein orf151 [imported] - Guillardia theta nucleomorph
C:Species: nucleomorph Guillardia theta
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: B90102
R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Rei
Nature 410, 1091-1096, 2001
A:Title: The highly reduced genome of an enslaved algal nucleus.
A:Reference number: A99082; MUID:11323671; PMID:11323671
A:Accession: B90102
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-151 <DOU>
A:Cross-references: UNIPROT:Q9AQY4; GB:AJ010592; NID:g12580657; PIDN:CAC26974.1; GSPDB:C

C:Genetics:

A:Gene: orf151

A:Map position: 2

A:Genome: nucleomorph

C:Keywords: nucleomorph

Alignment Scores:
Pred. No.: 33.4 Length: 151
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.60% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x B90102 (1-151)

QY 594 TCCTCACTGTTTCTCTTTTC 571

Db 28 SerSerLeuValPheLeuPhePhe 35

RESULT 11

F90101
hypothetical protein orf151b [imported] - Guillardia theta nucleomorph

C:Species: nucleomorph Guillardia theta
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: F90101
R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Rei
Nature 410, 1091-1096, 2001
A:Title: The highly reduced genome of an enslaved algal nucleus.
A:Reference number: A99082; MUID:11323671; PMID:11323671
A:Accession: F90101
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-151 <DOU>
A:Cross-references: UNIPROT:Q9AQY4; GB:AF165818; NID:g13794573; PIDN:AAK39948.1; GSPDB:C

C:Genetics:

A:Gene: orf151b

A:Map position: 1

A:Genome: nucleomorph

C:Keywords: nucleomorph

Alignment Scores:
Pred. No.: 33.4 Length: 151
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.60% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x F90101 (1-151)

QY 594 TCCTCACTGTTTCTCTTTTC 571

Db 28 SerSerLeuValPheLeuPhePhe 35

RESULT 12

C90118
hypothetical protein orf151 [imported] - Guillardia theta nucleomorph
C:Species: nucleomorph Guillardia theta
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: C90118
R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Rei
Nature 410, 1091-1096, 2001
A:Title: The highly reduced genome of an enslaved algal nucleus.
A:Reference number: A99082; MUID:11323671; PMID:11323671
A:Accession: C90118
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-151 <DOU>
A:Cross-references: UNIPROT:Q9AQY4; GB:AF083031; NID:g13794413; PIDN:AAK39790.1; GSPDB:C

C:Genetics:

A:Gene: orf151

A:Map position: 3

A:Genome: nucleomorph

C:Keywords: nucleomorph

Alignment Scores:
Pred. No.: 33.4 Length: 151
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.60% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x C90118 (1-151)

QY 594 TCCTCACTGTTTCTCTTTTC 571

Db 28 SerSerLeuValPheLeuPhePhe 35

RESULT 13

A90138
hypothetical protein orf151b [imported] - Guillardia theta nucleomorph
C:Species: nucleomorph Guillardia theta

A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: A90138
R;Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reid,
Nature 410, 1091-1096, 2001
A;Title: The highly reduced genome of an enslaved algal nucleus.
A;Reference number: A99082; MUID:11323671; PMID:11323671
A;Accession: A90138
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-151 <DOU>
A;Cross-references: UNIPROT:Q9A0Y4; GB:AF083031; NID:gl3794420; PIDN:AAK39797.1; GSPDB:G
C;Genetics:
A;Gene: orf15b
A;Map position: 3
A;Genome: nucleomorph
C;Keywords: nucleomorph

Alignment Scores:
Pred. No.: 33.4 Length: 151
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.60% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x A90138 (1-151)

QY 594 TCCTCACTGTGTTTCCTCTTTTC 571

Db 28 SerSerLeuValPheLeuPhePhe 35

RESULT 14

A59068

beta-casein variant CnH - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 20-Sep-1999 #sequence_revision 24-Sep-1999 #text_change 24-Sep-1999
C;Accession: A59068; B59068
R;Han, S.K.; Shin, Y.C.
Anim. Genet. 27(Suppl.2), 91b, 1996
A;Title: Biochemical characterization of the new beta-casein variant in Korean cattle.
A;Reference number: A59068
A;Accession: A59068
A;Status: protein sequence not shown
A;Molecule type: protein
A;Residues: 1-209 <HANI>
A;Experimental source: strain Korean cattle
A;Note: submitted to the Protein Sequence Database, September 1999
A;Note: includes casein phosphopeptide H
A;Accession: B59068
A;Status: protein sequence not shown
A;Molecule type: protein
A;Residues: 1-28 <HAN2>
A;Experimental source: strain Korean cattle
C;Superfamily: beta-casein
C;Keywords: milk; phosphoprotein
F;15,17,18,19/Binding site: phosphate (Ser) (covalent) #status predicted

Alignment Scores:
Pred. No.: 32.1 Length: 209
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.60% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x A59068 (1-209)

QY 470 TCCTTATCTTCTCAGAGAGTCT 447

Db 15 SerLeuSerSerSerGluGluSer 22

RESULT 15

JC1384

beta-casein precursor - goat
C;Species: Capra aegagrus hircus (domestic goat)
C;Date: 10-Jun-1993 #sequence_revision 10-Jun-1993 #text_change 09-Jul-2004
C;Accession: JC1384
R;Roberts, B.; DiTullio, P.; Vitale, J.; Hehir, K.; Gordon, K.
Gene 121, 255-262, 1992
A;Title: Cloning of the goat beta-casein-encoding gene and expression in transgenic mic
A;Reference number: JC1384; MUID:93077039; PMID:1446822
A;Accession: JC1384
A;Molecule type: DNA
A;Residues: 1-222 <ROB>
A;Cross-references: UNIPROT:P33048; GB:M90556
C;Genetics:
A;Gene: CSN2
A;Introns: 17/3; 26/3; 35/3; 45/3; 57/3; 221/3
C;Superfamily: beta-casein
C;Keywords: milk; phosphoprotein

Alignment Scores:
Pred. No.: 31.9 Length: 222
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.60% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x JC1384 (1-222)

QY 470 TCCTTATCTTCTCAGAGAGTCT 447

Db 30 SerLeuSerSerSerGluGluSer 37

Search completed: February 15, 2005, 02:30:27
Job time : 48 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 15, 2005, 02:04:41 ; Search time 168 Seconds
(without alignments)
4392.493 Million cell updates/sec

Title: AF334735
Perfect score: 307
Sequence: 1 TCGCCCTTCTCGCGCGCG.....AAAAAAAAAAAAAAAAAAAA 954

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3967864

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-Q=/cgn2_1/USPTO.spool_p/AF334735/runat_14022005_160237_19333/app_query.fasta_1.1095
-DB=A Geneseq 16Dec04 -Qfmt=fastan -SUFFIX=oligo.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=AF334735 @CGN 1.1.224 @runat_14022005_160237_19333 -NCPUs=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORE=0 -WAIT -DEPBLOCK=100 -LONGLOG
-DE TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : A Geneseq 16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	151	49.2	151	2 AAR79761	Aar79761 Human tes
2	151	49.2	151	2 AAW42642	Aaw42642 Human spe
3	151	49.2	162	2 AAW42692	Aaw42692 Recombina
4	76	24.8	119	3 AAG00833	Aag00833 Human sec
5	76	24.8	140	4 ABG11758	Abg11758 Novel hum
6	57	18.6	300	4 ABG08152	Abg08152 Novel hum
7	43	14.0	163	2 AAW42693	Aaw42693 Baboon sp
8	33	10.7	146	2 AAW42691	Aaw42691 Recombina
9	28	9.1	28	2 AAW42767	Aaw42767 Peptide o
10	25	8.1	25	2 AAW42762	Aaw42762 Peptide o

11	25	8.1	273	4 ABG11759	Abg11759 Novel hum
12	21	6.8	21	2 AAW42771	Aaw42771 Peptide d
13	21	6.8	21	2 AAW42770	Aaw42770 Peptide o
14	21	6.8	21	2 AAW42772	Aaw42772 Peptide d
15	19	6.2	19	2 AAW42869	Aaw42869 Peptide d
16	17	5.5	17	2 AAW42773	Aaw42773 Peptide d
17	17	5.5	18	2 AAW42874	Aaw42874 Antigenic
18	16	5.2	16	2 AAW42764	Aaw42764 Peptide o
19	16	5.2	16	2 AAW42868	Aaw42868 Peptide o
20	16	5.2	16	2 AAW42873	Aaw42873 Antigenic
21	13	4.2	13	2 AAW42870	Aaw42870 Sp17 pept
22	13	4.2	21	2 AAW42768	Aaw42768 Peptide o
23	13	4.2	35	2 AAW42871	Aaw42871 Sp17 pept
24	12	3.9	28	2 AAW42766	Aaw42766 Peptide o
25	11	3.6	16	2 AAW42763	Aaw42763 Peptide o
26	10	3.3	10	2 AAR79763	Aar79763 Rabbit te
27	10	3.3	10	2 AAR79764	Aar79764 Rabbit te
28	10	3.3	10	2 AAR79767	Aar79767 Rabbit te
29	10	3.3	10	2 AAR79777	Aar79777 Rabbit te
30	10	3.3	10	2 AAR79778	Aar79778 Rabbit te
31	10	3.3	10	2 AAR79766	Aar79766 Rabbit te
32	10	3.3	10	2 AAR79779	Aar79779 Rabbit te
33	10	3.3	10	2 AAR79773	Aar79773 Rabbit te
34	10	3.3	10	2 AAR79762	Aar79762 Rabbit te
35	10	3.3	10	2 AAR79772	Aar79772 Rabbit te
36	10	3.3	10	2 AAR79768	Aar79768 Rabbit te
37	10	3.3	10	2 AAR79770	Aar79770 Rabbit te
38	10	3.3	10	2 AAR79765	Aar79765 Rabbit te
39	10	3.3	10	2 AAR79769	Aar79769 Rabbit te
40	10	3.3	10	2 AAR79771	Aar79771 Rabbit te
41	10	3.3	10	2 AAW42733	Aaw42733 Antigenic
42	10	3.3	10	2 AAW42740	Aaw42740 Antigenic
43	10	3.3	10	2 AAW42744	Aaw42744 Antigenic
44	10	3.3	10	2 AAW42779	Aaw42779 Antigenic
45	10	3.3	10	2 AAW42793	Aaw42793 Antigenic

ALIGNMENTS

RESULT 1
AAR79761
ID AAR79761 standard; protein; 151 AA.
XX AC AAR79761;
XX 11-DEC-1995 (first entry)
XX Human testis sperm zona binding protein Sp17.
DE Human testis sperm zona binding protein Sp17.
XX Sperm antigen; autoantigen; zona pellucida; binding protein;
KW contraceptive; fertility.
XX Homo sapiens.
OS W09515764-A1.
PN 15-JUN-1995.
PD 17-NOV-1994; 94WO-US013328.
PF 10-DEC-1993; 93US-00166195.
PR (UYNC-) UNIV NORTH CAROLINA.
PA Orand MG, Widgren EE, Richardson RT, Lea IA;
XX WPI; 1995-224147/29.
XX N-PSDB; AAQ98173.
XX Novel sperm antigens useful as immunocontraceptive agents - and for
PT diagnosis of autoimmune infertility.
XX Claim 6; Page 30-31; 62pp; English.
PS

XX The cloning and sequence data of rabbit Sp17 was previously reported
 CC (Richardson and O'Rand, Mol. Biol. Cell 3, 15a (1992)). This protein is
 CC known to be a member of the rabbit sperm antigen (RSA) family of rabbit
 CC testis/sperm autoantigens and is also to be expressed in mice. A human
 CC testis cDNA library was screened using the protein coding region of the
 CC rabbit Sp17 gene as a probe. One clone contd. an insert 71% identical to
 CC the rabbit Sp17 gene at the nt. level (see AAQ98173). This clone encoded
 CC a protein (AAR79761) with a calc. mol. wt. of 17,534 Da, 76.7% identical
 CC to the rabbit Sp17 and 71.8% identical to the mouse Sp17 protein sequence
 XX
 XX Sequence 151 AA;

Alignment Scores:
 Pred. No.: 5.3e-143 Length: 151
 Score: 151.00 Matches: 151
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 49.19% Indels: 0
 DB: 2 Gaps: 0

AF334735 (1-954) x AAR79761 (1-151)

QY 136 ATGTCGATTCATTCCTCCACACCCACTACCGAATTCACAGGATTGGGAATCTTCT 195
 DB 1 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeuLeu 20
 QY 196 GAAGGCTACACGGCAGATCTCTGAGAGACACCGGCAATATACCAGCTTTTCAGCA 255
 DB 21 GluGlyLeuThrArgGluIleuArgGluGlnProAspAsnIleProAlaPheAla 40
 QY 256 GCTATTTTGGAGCGCTCTAGAGAAAGAGAGAAACCAACTTTTGATCCACAGAAATGG 315
 DB 41 AlaTyrPheGluSerLeuLeuGluLysArgGluLysThrAsnPheAspProAlaGluTrp 60
 QY 316 GGGAGTAAGTAGAGACCGCTTCTATACCAATCATGCTTCGAGAGCAAGAACCACT 375
 DB 61 GlySerLysValGluAspArgPheTyrAsnAsnHisAlaPheGluGluGlnProPro 80
 QY 376 GAGAAAGTGCCTTAACCAAGAGAGTCTCAGATATCTGGAGAGGAGAGAGAGATCA 435
 DB 81 GluLysSerAspProLysGlnGluSerGlnIleSerGlyLysGluGluGluThrSer 100
 QY 436 GTCACCATCTAGACTCTTCGGGAGACATAGCCAGAGAGGAGGCAAGAAATGAAAAA 495
 DB 101 ValThrIleLeuAspSerGluGluAspLysGluLysGluLysValAlaAlaValLys 120
 QY 496 ATCCAGCTGCCTTCGGGGACACATAGCCAGAGAGGAGGCAAGAAATGAAAAA 555
 DB 121 IleGlnAlaAlaPheArgGlyHisIleAlaArgGluGluAlaLysLysMetLysThrAsn 140
 QY 556 AGTCTTCAAAATGAGAAAGAGGAAACAG 588
 DB 141 SerLeuGlnAsnGluLysGluGluAsnLys 151

RESULT 2

AAW42642

ID AAW42642 standard; protein; 151 AA.

XX

XX AAW42642;

XX 27-APR-1998 (first entry)

XX

XX Human sperm zona binding protein Sp17.

XX

XX Sperm zona binding protein Sp17; zona pellucida; binding; reduction;
 KW Human sperm autoantigen; antigen; antigenic epitope; immunocontraception;
 KW fertility; immunocontraceptive vaccine; autoimmune fertility.

XX

OS Homo sapiens.

XX

PN W09739020-A2.

XX

PD 23-OCT-1997.
 XX
 XX 11-APR-1997; 97WO-US006489.
 XX
 XX 15-APR-1996; 96US-00632535.
 XX
 XX (UYN-C) UNIV NORTH CAROLINA.
 XX
 XX Orand MG, Lea I, Widgren EE;
 XX
 XX WPI; 1997-526394/48.
 DR
 DR N-PSDB; AAV04818.
 XX

Sperm protein Sp17 antigenic peptide(s) - used as immuno:contraceptive to
 reduce fertility.

Example 1; Page 59; 160pp; English.

The present sequence represents a human sperm zona binding protein Sp17.
 The cDNA sequence was identified by screening a human testis cDNA library
 with the coding region of the rabbit Sp17 gene as a probe. The human Sp17
 protein binds to zona pellucida at high affinity by binding sulphated,
 complex carbohydrates. The protein has been shown to be a human sperm
 autoantigen. Peptides comprising a contiguous segment of the amino acid
 of Sp17 have been found to be antigenic in several species. Peptides
 AAW42774-867 contain antigenic epitopes in mice, rabbits, non human
 primates and humans. The peptides can be used in immunocontraceptive
 methods to reduce the fertility of animals, particularly in females. The
 peptides can be used in an immunocontraceptive vaccine. The presence of
 the antigenic peptides can be detected to diagnose autoimmune fertility
 in both male and female subjects

Sequence 151 AA;

Alignment Scores:

Pred. No.: 5.3e-143 Length: 151
 Score: 151.00 Matches: 151
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 49.19% Indels: 0
 DB: 2 Gaps: 0

AF334735 (1-954) x AAW42642 (1-151)

QY 136 ATGTCGATTCATTCCTCCACACCCACTACCGAATTCACAGGATTGGGAATCTTCT 195
 DB 1 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeuLeu 20
 QY 196 GAAGGCTACACGGCAGATCTCTGAGAGACACCGGCAATATACCAGCTTTTCAGCA 255
 DB 21 GluGlyLeuThrArgGluIleuArgGluGlnProAspAsnIleProAlaPheAla 40
 QY 256 GCTATTTTGGAGCGCTTCTATACCAATCATGCTTCGAGAGCAAGAACCACTTTGATCCAGCA 315
 DB 41 AlaTyrPheGluSerLeuLeuGluLysArgGluLysThrAsnPheAspProAlaGluTrp 60
 QY 316 GGGAGTAAGTAGAGACCGCTTCTATACCAATCATGCTTCGAGAGCAAGAACCACT 375
 DB 61 GlySerLysValGluAspArgPheTyrAsnAsnHisAlaPheGluGluGlnProPro 80
 QY 376 GAGAAAGTGCCTTAACCAAGAGAGTCTCAGATATCTGGAGAGGAGAGAGACATCA 435
 DB 81 GluLysSerAspProLysGlnGluSerGlnIleSerGlyLysGluGluGluThrSer 100
 QY 436 GTCACCATCTAGACTCTTCTGAGAGAGATAGGAAAGAGAGAGAGGTTGCTGCTCAAA 495
 DB 101 ValThrIleLeuAspSerGluGluAspLysGluLysGluLysValAlaAlaValLys 120
 QY 496 ATCCAGCTGCCTTCGGGGACACATAGCCAGAGAGGAGGCAAGAAATGAAAAA 555
 DB 121 IleGlnAlaAlaPheArgGlyHisIleAlaArgGluGluAlaLysLysMetLysThrAsn 140
 QY 556 AGTCTTCAAAATGAGAAAGAGGAAACAG 588

```

Db      141 SerLeuGlnAsnGluGluLysGluGluAsnLys 151
|||||
RESULT 3
AAW42692
ID      AAW42692 standard; protein; 162 AA.
XX
AC      AAW42692;
XX
DT      27-APR-1998 (first entry)
XX
DE      Recombinant human Sp17 protein.
XX
KW      Sp17; human sperm antigen; sperm autoantigen; antigenic epitope;
KW      immunocontraction; fertility; autoimmune fertility; antigen;
KW      immunocontraceptive vaccine.
XX
OS      Synthetic.
OS      Homo sapiens.
XX
PN      WO9739020-A2.
XX
PD      23-OCT-1997.
XX
PF      11-APR-1997; 97WO-US006489.
XX
PR      15-APR-1996; 96US-00632535.
XX
PA      (UYNC-) UNIV NORTH CAROLINA.
XX
PI      Orand MG, Lea I, Widgren BE;
XX
PI      WPI; 1997-526394/48.
XX
PT      Sperm protein Sp17 antigenic peptide(s) - used as immuno:contraceptive to
PT      reduce fertility.
XX
PS      Example 8; Page 80; 160pp; English.
XX
CC      The present protein represents recombinant human Sp17 protein. The
CC      recombinant human Sp17 protein is expressed with an N-terminal containing
CC      the sequence Arg-Gly-Ser, followed by 6 His and Gly-Ser, all of which
CC      precede the Sp17 amino acids. Mice were immunised with the recombinant
CC      human Sp17 protein, and mated. A 42% decrease in pregnancy was observed
CC      with mice treated with the recombinant protein. Peptides comprising a
CC      contiguous segment of the amino acids of Sp17 have been also been found
CC      to be antigenic. They induce antibodies which recognise sperm, and
CC      inhibit fertilisation. The peptides can be used in immunocontraceptive
CC      methods as a immunocontraceptive vaccine to reduce the fertility of
CC      animals, particularly in females. The presence of the antigenic peptides
CC      can be detected to diagnose autoimmune fertility in both male and female
CC      subjects
XX
SQ      Sequence 162 AA;

Alignment Scores:
Pred. No.: 5,25e-143 Length: 162
Score: 151.00 Matches: 151
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.19% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x AAW42692 (1-162)
QY      136 ATGTGATTCATTCCTCCACACCCACTACCGAATTCACAGATTGGGAATCTTCTT 195
Db      12 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeu 31
QY      196 GAAGGCTGACGCGGAGATTCTGAGAGCGACCGACATATACACCTTTTCGACGA 255
Db      32 GluGlyLeuThrArgGluIleLeuArgGluGlnProAspAsnIleProAlaPheAla 51

```

```

XX SQ Sequence 119 AA;
Alignment Scores:
Pred. No.: 3,09e-67 Length: 119
Score: 76.00 Matches: 76
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.76% Indels: 0
DB: Gaps: 0
AF334735 (1-954) x AAG00833 (1-119)
QY 136 ATGTCGATTCATTCTCCACACCCACTACCGAATTCACAGGATTTCGGATCTTCTT 195
DB 1 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeu 20
QY 196 GAAGGCTGACACGCGAGATTCTGAGAGAGCAACCGCAATATACACGCTTTTCGACGA 255
DB 21 GluGlyLeuThrArgGluLeuArgGluGlnProAspAsnIleProAlaPheAla 40
QY 256 GCCTATTTTTCAGAGCCTTCTAGAGAAAGAGAGAAACCAACTTTGATCCAGCAGATGG 315
DB 41 AlaTyrPheGluSerLeuLeuGluLysArgGluLysThrAsnPheAspProAlaGluTrp 60
QY 316 GCGAGTAAAGTAGAGACCGCTTCTATACAAATCATGCAATTCGAGGAG 363
DB 61 GlySerLysValGluAspArgPheTyrAsnAsnHisAlaPheGluGlu 76
RESULT 5
ABG11758
ID ABG11758 standard; protein; 140 AA.
XX
AC ABG11758;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #11749.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
XX
DR N-PSDB; AAS75945.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 42117; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal

```

```

CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 140 AA;
Alignment Scores:
Pred. No.: 3,03e-67 Length: 140
Score: 76.00 Matches: 76
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.76% Indels: 0
DB: Gaps: 0
AF334735 (1-954) x ABG11758 (1-140)
QY 310 GAATGGGGAGTAAGTAGAAGACCGCTTCTATACAAATCATGCTTCGAGGAGCAAGAA 369
DB 65 GluTrpGlySerLysValGluAspArgPheTyrAsnAsnHisAlaPheGluGluGlu 84
QY 370 CCACCTGAGAAAAGTGATCCTTAACAAGAAGTCTCAGATATCTGGGAGGAGGAGAG 429
DB 85 ProProGluLysSerAspProLysGlnGluSerGlnIleSerGlyLysGluGlu 104
QY 430 ACATCAGTCACCATCTTAGACTCTTCTGAGGAAGATAAGGAAAAAGAGAGGTTGCTGCT 489
DB 105 ThrSerValThrIleLeuAspSerSerGluGluAspLysGluLysGluGluValAla 124
QY 490 GTCAAAATCCCAAGTCGCTTCGGGGACACATAGCCAGAGAGGAGCA 537
DB 125 ValLysIleGlnAlaAlaPheArgGlyHisIleAlaArgGluGluAla 140
RESULT 6
ABG08152
ID ABG08152 standard; protein; 300 AA.
XX
AC ABG08152;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #8143.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.

```

DR N-PSDB; AAS72339.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 20; SEQ ID NO 39511; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIFO at
CC ftp.wifo.int/pub/published_pct_sequences
XX
XX Sequence 300 AA;

Alignment Scores:
Pred. No.: 4,27e-48 Length: 300
Score: 57.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.57% Indels: 0
DB: 4 Gaps: 0

AF334735 (1-954) x ABG08152 (1-300)

QY 118 GCAGTTCTTACCAAGAGATGTCGATTCCATTCTCCACACCCACTACCGAATTCACAA 177
|||
Db 223 AlaValLeuThrIlyslsMetSerIleProPheSerAsnThrHisTyArgIleProGln 242
|||
QY 178 GGAATTTGGGAATCTTTGAGGGCTGACACGCGAGATCTGAGAGACGACCGGCAAT 237
|||
Db 243 GlyPheGlyAsnLeuLeuGluGlyLeuThrArgGluIleLeuArgGluGlnProAspAsn 262
|||
QY 238 ATACCGAGCTTTTCAGCAGCCTATTTTGGAGCGCTTCTAGAGAAAGAGAG 288
|||
Db 263 IleProAlaPheAlaAlaIleTyPheGluSerLeuLeuGluIlyslsArgGlu 279
|||

RESULT 7

AAW42693
ID AAW42693 standard; protein; 163 AA.

AC AAW42693;

DT 27-APR-1998 (first entry)

DE Baboon sperm zona binding protein Sp17 (BSp17).

XX Sperm zona binding protein Sp17; zona pellucida; binding; reduction;
KW Human sperm autoantigen; antigen; antigenic epitope; immunoreception;
KW fertility; immunoreceptive vaccine; autoimmune fertility.

OS Papio sp.

XX WO9739020-A2.

XX

PD 23-OCT-1997.
XX
XX 11-APR-1997; 97WO-US006489.
XX
XX 15-APR-1996; 96US-00632535.
XX
XX (UYNC-) UNIV NORTH CAROLINA.
XX
XX Orand MG, Lea I, Widgren EE;
XX
XX WPI; 1997-526394/48.
XX N-PSDB; AAV04821.
XX
XX Sperm protein Sp17 antigenic peptide(s) - used as immuno:contraceptive to
PT reduce fertility.
XX
XX Example 11; Fig 16A; 160pp; English.
XX
XX The present sequence represents a baboon sperm zona binding protein Sp17.
CC The CDNA sequence was identified by screening a baboon testis cDNA
CC library with the coding region of the human Sp17 gene as a probe. The
CC human Sp17 protein binds to zona pellucida at high affinity by binding
CC sulphated, complex carbohydrates. Peptides comprising a contiguous
CC segment of the amino acids of human Sp17 have been found to be antigenic
CC in several species. Peptides AAW42774-867 contain antigenic epitopes in
CC mice, rabbits, non human primates and humans. The peptides can be used in
CC immunocontraceptive methods to reduce the fertility of animals,
CC particularly in females. The peptides can be used in an
CC immunocontraceptive vaccine. The presence of the antigenic peptides can
CC be detected to diagnose autoimmune fertility in both male and female
CC subjects
XX
XX Sequence 163 AA;

Alignment Scores:
Pred. No.: 6,38e-34 Length: 163
Score: 43.00 Matches: 70
Percent Similarity: 95.89% Conservative: 0
Best Local Similarity: 95.89% Mismatches: 0
Query Match: 14.01% Indels: 3
DB: 2 Gaps: 0

AF334735 (1-954) x AAW42693 (1-163)

QY 136 ATGTCGATTCCATTCTCCACACCCACTACCGAATTCACAAAGGATTTGGGAATCTCTT 195
|||
Db 1 MetSerIleProPheSerAsnThrHisTyArgIleProGlnGlyPheIlyslsLeu 20
|||
QY 196 GAAGGGCTGACACGCGAGATCTGAGAGACGACCGGCAATATATACCGCTTTTGCAGCA 255
|||
Db 21 GluGlyLeuThrArgGluIleLeuArgGluGlnProAspAsnIleProAlaPhe---Ala 39
|||
QY 256 GCCTATTTTGGAGCGCTTCTAGAGAAAGAGAGAAACCACTTTGATCCAGCAGATGG 315
|||
Db 40 AlaTyPheGluSerLeuLeuGluIlyslsArgGluIlyslsThrAsnPheAspProAlaGluTrp 59
|||
QY 316 GGGAGTAGTAGAGACCGCTTCTATACCAAT 348
|||
Db 60 GlySerLysValGluAspArgPheTyrrAsnAsn 70
|||

RESULT 8

AAW42691
ID AAW42691 standard; protein; 146 AA.

AC AAW42691;

DT 27-APR-1998 (first entry)

DE Recombinant rabbit Sp17 protein.

XX Sp17; rabbit sperm antigen; RSA; sperm autoantigen; antigen;

KW antigenic epitope; immunoreception; fertility; autoimmune fertility;

KW immunocontraceptive vaccine.

XX Synthetic.
 OS Oryctolagus cuniculus.
 XX WO9739020-A2.
 PN 23-OCT-1997.
 XX 11-APR-1997; 97WO-US006489.
 XX 15-APR-1996; 96US-00632535.
 XX (UYN-) UNIV NORTH CAROLINA.
 XX Orand MG, Lea I, Widgren EE;
 XX WPI; 1997-526394/48.
 XX Sperm protein Sp17 antigenic peptide(s) - used as immuno:contraceptive to reduce fertility.
 XX Example 6; Page 79; 160pp; English.
 XX The present protein represents rabbit recombinant Sp17 protein. The recombinant rabbit Sp17 protein is expressed minus the first 11 N-terminal amino acids, but with an N-terminal containing the sequence Arg-Gly-Ser, followed by 6 His and Gly-Ser, all of which precede the Sp17 amino acids. The binding of of immune sera from a female rabbit immunised with the recombinant fusion Sp17 protein to the Sp17 sequential peptides AAW42643-89 was determined. Peptides comprising a contiguous segment of the amino acid of Sp17 have been found to be antigenic. They induce antibodies which recognise sperm, and inhibit fertilisation. The peptides can be used in immunocontraceptive methods as a immunocontraceptive vaccine to reduce the fertility of animals, particularly in females. The presence of the antigenic peptides can be detected to diagnose autoimmune fertility in both male and female subjects

XX SQ Sequence 146 AA;
 Alignment Scores:
 Pred. No.: 8.15e-24 Length: 146
 Score: 33.00 Matches: 33
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 10.75% Indels: 0
 DB: 2 Gaps: 0

AF334735 (1-954) x AAW42691 (1-146)
 QY 169 ATTCCACAAAGATTGGGAATCTTTGAAGGCTGACACGGAGATTCTGAGAGACAA 228
 Db 12 IleProGInGlyPheGlyAsnLeuLeuGluGlyLeuThrArgGluIleLeuArgGluIn 31
 QY 229 CCGGACAATATACCACTTTTCAGCAGCAGCTATTGTAG 267
 Db 32 ProAspAsnIleProAlaPheAlaAlaLaTyPheGlu 44

RESULT 9
 AAW42767
 ID AAW42767 standard; peptide; 28 AA.
 XX AAW42767;
 XX 27-APR-1998 (first entry)
 XX Peptide of the specification.
 XX Sp17; sperm autoantigen; antigen; antigenic epitope; fertility;
 XX immunocontraception; immunocontraceptive vaccine; autoimmune fertility.
 XX Synthetic.
 XX WO9739020-A2.
 PN 23-OCT-1997.

XX 23-OCT-1997.
 XX 11-APR-1997; 97WO-US006489.
 XX 15-APR-1996; 96US-00632535.
 XX (UYN-) UNIV NORTH CAROLINA.
 XX Orand MG, Lea I, Widgren EE;
 XX WPI; 1997-526394/48.
 XX Sperm protein Sp17 antigenic peptide(s) - used as immuno:contraceptive to reduce fertility.
 XX Disclosure; Page 77; 160pp; English.
 XX Peptides AAW42762-70 appear in the specification. A series of peptides derived from rabbit and human Sp17 proteins are also disclosed. These Sp17 peptides contain antigenic epitopes in mice, rabbits, non-human primates and humans. They induce antibodies which recognise sperm, and inhibit fertilisation. The peptides can be used in immunocontraceptive methods as a immunocontraceptive vaccine to reduce the fertility of animals, particularly in females. The presence of the antigenic peptides can be detected to diagnose autoimmune fertility in both male and female subjects

XX SQ Sequence 28 AA;
 Alignment Scores:
 Pred. No.: 1.13e-18 Length: 28
 Score: 28.00 Matches: 28
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 9.12% Indels: 0
 DB: 2 Gaps: 0

AF334735 (1-954) x AAW42767 (1-28)
 QY 298 TTGTATCCAGCAGATGGGAGTAGGTAGAGCCGCTTCTATACATCATCATTC 357
 Db 1 PheAspProAlaGluTrpGlySerTysValGluAspArgPheTyAsnAsnHisAlaPhe 20
 QY 358 GAGGAGCAAGAACCACTGAGAAA 381
 Db 21 GluGluGlnGluProProGluLys 28

RESULT 10
 AAW42762
 ID AAW42762 standard; peptide; 25 AA.
 XX AAW42762;
 XX 27-APR-1998 (first entry)
 XX Peptide of the specification.
 XX Sp17; sperm autoantigen; antigen; antigenic epitope; fertility;
 XX immunocontraception; immunocontraceptive vaccine; autoimmune fertility.
 XX Synthetic.
 XX WO9739020-A2.
 XX 23-OCT-1997.
 XX 11-APR-1997; 97WO-US006489.
 XX 15-APR-1996; 96US-00632535.
 XX (UYN-) UNIV NORTH CAROLINA.
 XX

PI Orand MG, Lea I, Widgren EB;
XX WPI; 1997-526394/48.
XX Sperm protein Sp17 antigenic peptide(s) - used as immuno:contraceptive to
PT reduce fertility.
PT Disclosure; Page 75; 160pp; English.
XX Peptides AAW42762-70 appear in the specification. A series of peptides
CC derived from rabbit and human Sp17 proteins are also disclosed. These
CC Sp17 peptides contain antigenic epitopes in mice, rabbits, non-human
CC primates and humans. They induce antibodies which recognise sperm, and
CC inhibit fertilisation. The peptides can be used in immunocontraceptive
CC methods as a immunocontraceptive vaccine to reduce the fertility of
CC animals, particularly in females. The presence of the antigenic peptides
CC can be detected to diagnose autoimmune fertility in both male and female
CC subjects
XX
SQ Sequence 25 AA;

Alignment Scores:
Pred. No.: 1-22e-15 Length: 25
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.14% Indels: 0
DB: Gaps: 0

AF334735 (1-954) x AAW42762 (1-25)
QY 145 CCATTCTCCACACCCACTACCGAATTCACAGGATTTGGGAATCTTTCGAAGGCTG 204
Db 1 ProPheSerAsnThrHisTyArgIleProGlnGlyPheGlyAsnLeuLeuGluGlyLeu 20
QY 205 ACACGGCAGATTCTG 219
Db 21 ThrArgGluIleLeu 25

RESULT 11
ABG11759
ID ABG11759 standard; protein; 273 AA.
XX AC ABG11759;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #11750.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
XX OS WO200175067-A2.
XX PN 11-OCT-2001.
XX PP 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX DR N-PSDB; AAS75946.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess

PT biodiversity.
XX Claim 20; SEQ ID NO 42118; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 273 AA;

Alignment Scores:
Pred. No.: 9.07e-16 Length: 273
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.14% Indels: 0
DB: Gaps: 0

AF334735 (1-954) x ABG11759 (1-273)
QY 436 GTCACATCTTAGACTCTTCTGAGGAGATAGGAAAGAGAGGTTCTGCTGTCAA 495
Db 62 ValThrIleLeuAspSerGluGluAspLysGluLysGluLysValAlaValLys 81
QY 496 ATCCAAGCTGCCTTC 510
Db 82 IleGlnAlaAlaPhe 86

RESULT 12
AAW42771
ID AAW42771 standard; peptide; 21 AA.
XX AC AAW42771;
XX DT 27-APR-1998 (first entry)
XX DE Peptide derived from Sp17.
XX KW Sp17; sperm autoantigen; antigen; antigenic epitope; fertility;
KW immunocontraception; immunocontraceptive vaccine; autoimmune fertility.
XX OS Synthetic.
XX PN WO9739020-A2.
XX PD 23-OCT-1997.
XX PF 11-APR-1997; 97WO-US006489.
XX PR 15-APR-1996; 96US-00632535.
XX PA (UYNC-) UNIV NORTH CAROLINA.
XX Orand MG, Lea I, Widgren EB;
XX

```
DR WPI; 1997-526394/48.
XX
PT Sperm protein Sp17 antigenic peptide(s) - used as immuno:contraceptive to
PT reduce fertility.
XX
PS Disclosure; Page 83; 160pp; English.
XX
CC The present peptide is derived from a Sp17 protein. Sp17 peptides contain
CC antigenic epitopes in mice, rabbits, non-human primates and humans. They
CC induce antibodies which recognise sperm, and inhibit fertilisation. The
CC peptides can be used in immunocontraceptive methods as a
CC immunocontraceptive vaccine to reduce the fertility of animals,
CC particularly in females. The presence of the antigenic peptides can be
CC detected to diagnose autoimmune fertility in both male and female
CC subjects
XX
SQ Sequence 21 AA;
Alignment Scores:
Pred. No.: 1.37e-11 Length: 21
Score: 21.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.84% Indels: 0
DB: 2 Gaps: 0
AF334735 (1-954) x AAW42771 (1-21)
QY 166 CGAATTCACAGAGATTTCGGATCTTGAAGGCTGACACGCGAGATTCTGAGAGAG 225
Db 1 ArgileProGlnGlyPheGlyAsnLeuLeuGluGlyLeuThrArgGluIleLeuArgGlu 20
QY 226 CAA 228
Db 21 Gln 21
RESULT 13
AAW42770
ID AAW42770 standard; peptide; 21 AA.
XX
AC AAW42770;
XX
DT 27-APR-1998 (first entry)
XX
DE Peptide of the specification.
XX
KW Sp17; sperm autoantigen; antigen; antigenic epitope; fertility;
KW immunocontraception; immunocontraceptive vaccine; autoimmune fertility.
XX
OS Synthetic.
XX
PN WO9739020-A2.
XX
PD 23-OCT-1997.
XX
PF 11-APR-1997; 97WO-US006489.
XX
PR 15-APR-1996; 96US-00632535.
XX
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Orand MG, Lea I, Widgren EE;
XX
WPI; 1997-526394/48.
XX
Sperm protein Sp17 antigenic peptide(s) - used as immuno:contraceptive to
XX reduce fertility.
XX
PS Disclosure; Page 78; 160pp; English.
XX
CC Peptides AAW42762-70 appear in the specification. A series of peptides
CC derived from rabbit and human Sp17 proteins are also disclosed. These
CC Sp17 peptides contain antigenic epitopes in mice, rabbits, non-human
CC
```

```
CC primates and humans. They induce antibodies which recognise sperm, and
CC inhibit fertilisation. The peptides can be used in immunocontraceptive
CC methods as a immunocontraceptive vaccine to reduce the fertility of
CC animals, particularly in females. The presence of the antigenic peptides
CC can be detected to diagnose autoimmune fertility in both male and female
CC subjects
XX
SQ Sequence 21 AA;
Alignment Scores:
Pred. No.: 1.37e-11 Length: 21
Score: 21.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.84% Indels: 0
DB: 2 Gaps: 0
AF334735 (1-954) x AAW42770 (1-21)
QY 481 GTTGCTGCTGTCAAAATCCAAAGCTGCTTCGGGACACATAGCCAGAGAGGAGCAAG 540
Db 1 ValAlaAlaVallylleGlnAlaAlaPheArgGlyHisIleAlaArgGluGluAlaLys 20
QY 541 AAA 543
Db 21 Lys 21
RESULT 14
AAW42772
ID AAW42772 standard; peptide; 21 AA.
XX
AC AAW42772;
XX
DT 27-APR-1998 (first entry)
XX
DE Peptide derived from Sp17.
XX
KW Sp17; sperm autoantigen; antigen; antigenic epitope; fertility;
KW immunocontraception; immunocontraceptive vaccine; autoimmune fertility.
XX
OS Synthetic.
XX
PN WO9739020-A2.
XX
PD 23-OCT-1997.
XX
PF 11-APR-1997; 97WO-US006489.
XX
PR 15-APR-1996; 96US-00632535.
XX
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Orand MG, Lea I, Widgren EE;
XX
WPI; 1997-526394/48.
XX
Sperm protein Sp17 antigenic peptide(s) - used as immuno:contraceptive to
XX reduce fertility.
XX
PS Claim 1; Page 83; 160pp; English.
XX
CC The present peptide is derived from a Sp17 protein. Sp17 peptides contain
CC antigenic epitopes in mice, rabbits, non-human primates and humans. They
CC induce antibodies which recognise sperm, and inhibit fertilisation. The
CC peptides can be used in immunocontraceptive methods as a
CC immunocontraceptive vaccine to reduce the fertility of animals,
CC particularly in females. The presence of the antigenic peptides can be
CC detected to diagnose autoimmune fertility in both male and female
CC subjects
XX
SQ Sequence 21 AA;
Alignment Scores:
```

Pred. No.: 1.37e-11 Length: 21
Score: 21.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.84% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x AAW42772 (1-21)

QY 319 AGTAAGGTAGAGACCGCTTCTATACATCATGCGAGGAGCAAGAACCACTGAG 378
Db 1 SerLysValGluAspArgPheTyrAsnAsnHisAlaPheGluGluGluProGlu 20

QY 379 AAA 381
Db 21 Lys 21

RESULT 15

AAW42869
ID AAW42869 standard; peptide; 19 AA.

XX AC AAW42869;

DT 27-APR-1998 (first entry)

DE Peptide of the specification.

KW Sp17; sperm autoantigen; antigen; antigenic epitope; fertility;
XX immunocontraception; immunocontraceptive vaccine; autoimmune fertility.

OS Synthetic.

PN WO9739020-A2.

PD 23-OCT-1997.

PF 11-APR-1997; 97WO-US006489.

PR 15-APR-1996; 96US-00632535.

PA (UYN-) UNIV NORTH CAROLINA.

PI Orand MG, Lea I, Widgren BE;

DR WPI; 1997-526394/48.

PT Sperm protein Sp17 antigenic peptide(s) - used as immuno:contraceptive to
XX reduce fertility.

PS Disclosure; Page 134; 160pp; English.

XX The present peptide appears in the specification. A series of peptides
CC derived from rabbit and human Sp17 proteins are also disclosed. These
CC Sp17 peptides contain antigenic epitopes in mice, rabbits, non-human
CC primates and humans. They induce antibodies which recognise sperm, and
CC inhibit fertilisation. The peptides can be used in immunocontraceptive
CC methods as a immunocontraceptive vaccine to reduce the fertility of
CC animals, particularly in females. The presence of the antigenic peptides
CC can be detected to diagnose autoimmune fertility in both male and female
CC subjects

SQ Sequence 19 AA;

Alignment Scores:

Pred. No.: 1.46e-09 Length: 19
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.19% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x AAW42869 (1-19)

QY 487 GCTGTCAAATCCAAGCTGCTTCCGGGACACATAGCCAGAGAGGAGCAAGAAA 543
Db 1 AlaValLysIleGlnAlaAlaPheArgGlyHisIleAlaArgGluGluAlaLysLys 19

Search completed: February 15, 2005, 02:22:02
Job time : 171 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 15, 2005, 02:14:06 ; Search time 37.5 Seconds
(without alignments)
3798.144 Million cell updates/sec

Title: AF334735
Perfect score: 307
Sequence: 1 TCGCCCTCTCGCGCGCG.....AAAAAAAAAAAAAAAAAAAA 954

Scoring table:
OLIGO 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delopt 6.0, Delext 7.0
Searched: 513545 seqs, 74649064 residues

Word size: 1
Total number of hits satisfying chosen parameters: 903960

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q/cgn2_1/USPTO_spool_P/AF334735/runat_14022005_160239_19363/app_query.fasta_1.1095
-DB=issued_Patents_AA -OPT=fastan -SUFFIX=oligo.rai -MINMATCH=0.1 -LOOPCL=70
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=AF334735_QCGN 1.1.46 @runat_14022005_160239_19363 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -XGAPOP=6
-XGAPEXT=7 -XGAPOP=60 -XGAPEXT=6 -DELEXT=7

Database : Issued Patents AA:
1: /cgn2_6/prodata/1/iaa/5A COMB pep.*
2: /cgn2_6/prodata/1/iaa/5B COMB pep.*
3: /cgn2_6/prodata/1/iaa/6A COMB pep.*
4: /cgn2_6/prodata/1/iaa/6B COMB pep.*
5: /cgn2_6/prodata/1/iaa/PCTUS COMB pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	151	49.2	151	1	US-08-166-195A-2
2	151	49.2	151	2	US-08-436-772-2
3	151	49.2	151	4	US-08-436-883B-2
4	76	24.8	119	4	US-09-513-999C-4914
5	44	14.3	145	1	US-08-166-195A-50
6	44	14.3	146	2	US-08-436-772-50
7	44	14.3	146	2	US-08-436-883B-50
8	44	14.3	148	1	US-08-166-195A-51
9	44	14.3	148	2	US-08-436-772-51
10	44	14.3	148	2	US-08-436-883B-51
11	28	9.1	28	2	US-08-436-772-57
12	28	9.1	28	2	US-08-436-883B-57

13	25	8.1	25	2	US-08-436-772-52	Sequence 52, Appl
14	25	8.1	25	2	US-08-436-883B-52	Sequence 52, Appl
15	21	6.8	21	2	US-08-436-772-60	Sequence 60, Appl
16	21	6.8	21	2	US-08-436-883B-60	Sequence 60, Appl
17	16	5.2	16	2	US-08-436-772-54	Sequence 54, Appl
18	16	5.2	16	2	US-08-436-883B-54	Sequence 54, Appl
19	13	4.2	21	2	US-08-436-772-58	Sequence 58, Appl
20	13	4.2	21	2	US-08-436-883B-58	Sequence 58, Appl
21	12	3.9	28	2	US-08-436-772-56	Sequence 56, Appl
22	12	3.9	28	2	US-08-436-883B-56	Sequence 56, Appl
23	11	3.6	16	2	US-08-436-772-53	Sequence 53, Appl
24	11	3.6	16	2	US-08-436-883B-53	Sequence 53, Appl
25	10	3.3	10	1	US-08-166-195A-3	Sequence 3, Appl
26	10	3.3	10	1	US-08-166-195A-4	Sequence 4, Appl
27	10	3.3	10	1	US-08-166-195A-5	Sequence 5, Appl
28	10	3.3	10	1	US-08-166-195A-6	Sequence 6, Appl
29	10	3.3	10	1	US-08-166-195A-7	Sequence 7, Appl
30	10	3.3	10	1	US-08-166-195A-8	Sequence 8, Appl
31	10	3.3	10	1	US-08-166-195A-9	Sequence 9, Appl
32	10	3.3	10	1	US-08-166-195A-10	Sequence 10, Appl
33	10	3.3	10	1	US-08-166-195A-11	Sequence 11, Appl
34	10	3.3	10	1	US-08-166-195A-12	Sequence 12, Appl
35	10	3.3	10	1	US-08-166-195A-13	Sequence 13, Appl
36	10	3.3	10	1	US-08-166-195A-14	Sequence 14, Appl
37	10	3.3	10	1	US-08-166-195A-18	Sequence 18, Appl
38	10	3.3	10	1	US-08-166-195A-19	Sequence 19, Appl
39	10	3.3	10	1	US-08-166-195A-20	Sequence 20, Appl
40	10	3.3	10	2	US-08-436-772-3	Sequence 3, Appl
41	10	3.3	10	2	US-08-436-772-4	Sequence 4, Appl
42	10	3.3	10	2	US-08-436-772-5	Sequence 5, Appl
43	10	3.3	10	2	US-08-436-772-6	Sequence 6, Appl
44	10	3.3	10	2	US-08-436-772-7	Sequence 7, Appl
45	10	3.3	10	2	US-08-436-772-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-08-166-195A-2
; Sequence 2, Application US/08166195A
; Patent No. 5480799
; GENERAL INFORMATION:
; APPLICANT: O'Rand, Michael G.
; APPLICANT: Widgren, Esther E.
; APPLICANT: Richardson, Richard T.
; APPLICANT: Lea, Isabel
; TITLE OF INVENTION: Sperm Antigen Corresponding to a
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenneth D. Sibley
; STREET: P. O. Box 34009
; CITY: Charlotte
; STATE: No. 5480799th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/166,195A
; FILING DATE: 10 DEC 1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470/73
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 2:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 151 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-166-195A-2

Alignment Scores:
Pred. No.: 2,346-135 Length: 151
Score: 151.00 Matches: 151
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.19% Indels: 0
DB: 1 Gaps: 0

AF334735 (1-954) x US-08-166-195A-2 (1-151)

QY 136 ATGTCGATTCCATTCTCCACACCCACTACCGAATCCACAGGATTTGGGAATCTTCTT 195
Db 1 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeuLeu 20
QY 196 GAAGGCTGACGCGAGATTCTGAGAGACCAACCGGACAATATACCACTTTTGCAGCA 255
Db 21 GluGlyLeuThrArgGluIleLeuArgGluGlnProAspAsnIleProAlaPheAlaAla 40
QY 256 GCCTATTCTTGGAGCCCTCTAGAGAAAGAGAGAAACCAACTTTGATCCAGCAATGG 315
Db 41 AlaTyrPheGluSerLeuLeuGluLysArgGluLysThrAsnPheAspProAlaGluTrp 60
QY 316 GGGAGTAAGTAGAAGACCGCTTCTATAACAATCATGCAATTCGAGGAGCAAGACCACT 375
Db 61 GlySerLysValGluAspArgPheTyrAsnAsnHisAlaPheGluGluGlnProPro 80
QY 376 GAGAAAGTGTCTTAAACAAGAGAGTCTCAGATATCTGGAGAGAGAGAGAGAGATCA 435
Db 81 GluLysSerAspProLysGlnGluSerGlnLysThrAsnPheAspProAlaGluTrp 100
QY 436 GTCACCATCTTAGACTCTTCTGAGGAAGATAGGAAAGAGAGAGTTTGTCTGCTCAAA 495
Db 101 ValThrIleLeuAspSerSerGluGluAspLysGluLysGluGluValAlaAlaVallys 120
QY 496 ATCCAAGCTGCCTTCGCGGGACACATAGCAGAGAGGCGGCAAGAAATGAAACAAAT 555
Db 121 IleGlnAlaAlaPheArgGlyHisIleAlaArgGluGluAlaLysMetLysThrAsn 140
QY 556 AGTCTTCAAAATGAGGAAAAAGGAAAAACAAG 588
Db 141 SerLeuGlnAsnGluGluLysGluGluAsnLys 151

RESULT 3
US-08-436-772-2
; Sequence 2, Application US/08436883B
; Patent No. 5820861
; GENERAL INFORMATION:
; APPLICANT: O'Rand, Michael G.
; APPLICANT: Widgren, Esther E.
; APPLICANT: Richardson, Richard T.
; APPLICANT: Lea, Isabel
; TITLE OF INVENTION: Sperm Antigen Corresponding to a Sperm
; TITLE OF INVENTION: Zona Binding Protein Autoantigenic Epitope
; NUMBER OF SEQUENCES: 60

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,772
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-73B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 151 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-436-772-2

Alignment Scores:
Pred. No.: 2,346-135 Length: 151
Score: 151.00 Matches: 151
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.19% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x US-08-436-772-2 (1-151)

QY 136 ATGTCGATTCCATTCTCCACACCCACTACCGAATCCACAGGATTTGGGAATCTTCTT 195
Db 1 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeuLeu 20
QY 196 GAAGGCTGACGCGAGATTCTGAGAGACCAACCGGACAATATACCACTTTTGCAGCA 255
Db 21 GluGlyLeuThrArgGluIleLeuArgGluGlnProAspAsnIleProAlaPheAlaAla 40
QY 256 GCCTATTCTTGGAGCCCTCTAGAGAAAGAGAGAAACCAACTTTGATCCAGCAATGG 315
Db 41 AlaTyrPheGluSerLeuLeuGluLysArgGluLysThrAsnPheAspProAlaGluTrp 60
QY 316 GGGAGTAAGTAGAAGACCGCTTCTATAACAATCATGCAATTCGAGGAGCAAGACCACT 375
Db 61 GlySerLysValGluAspArgPheTyrAsnAsnHisAlaPheGluGluGlnProPro 80
QY 376 GAGAAAGTGTCTTAAACAAGAGAGTCTCAGATATCTGGAGAGAGAGAGAGATCA 435
Db 81 GluLysSerAspProLysGlnGluSerGlnLysThrAsnPheAspProAlaGluTrp 100
QY 436 GTCACCATCTTAGACTCTTCTGAGGAAGATAGGAAAGAGAGAGTTTGTCTGCTCAAA 495
Db 101 ValThrIleLeuAspSerSerGluGluAspLysGluLysGluGluValAlaAlaVallys 120
QY 496 ATCCAAGCTGCCTTCGCGGGACACATAGCAGAGAGGCGGCAAGAAATGAAACAAAT 555
Db 121 IleGlnAlaAlaPheArgGlyHisIleAlaArgGluGluGluAlaLysMetLysThrAsn 140
QY 556 AGTCTTCAAAATGAGGAAAAAGGAAAAACAAG 588
Db 141 SerLeuGlnAsnGluGluLysGluGluAsnLys 151

RESULT 2
US-08-436-772-2
; Sequence 2, Application US/08436772
; Patent No. 5814456
; GENERAL INFORMATION:
; APPLICANT: O'Rand, Michael G.
; APPLICANT: Widgren, Esther E.
; APPLICANT: Richardson, Richard T.
; APPLICANT: Lea, Isabel
; TITLE OF INVENTION: Sperm Antigen Corresponding to a Sperm
; TITLE OF INVENTION: Zona Binding Protein Autoantigenic Epitope
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenneth D. Sibley
; STREET: P.O. Box 34009
; CITY: Charlotte
; STATE: No. 5814456th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Kenneth D. Sibley
;; STREET: P.O. Box 34009
;; CITY: Charlotte
;; STATE: No. 5820861th Carolina
;; COUNTRY: USA
;; ZIP: 28234
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/436,883B
;; FILING DATE: 08-MAY-1995
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Sibley, Kenneth D.
;; REGISTRATION NUMBER: 31,665
;; REFERENCE/DOCKET NUMBER: 5470-73C
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 919-420-2200
;; TELEFAX: 919-881-3175
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 151 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-436-883B-2

Alignment Scores:
Pred. No.: 2,34e-135 Length: 151
Score: 151.00 Matches: 151
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.19% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x US-08-436-883B-2 (1-151)

QY 136 ATGTCGATTCATCTCCACACCCACTACCGAATCCACAGATTTGGGAATCTTCTT 195
Db 1 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeu 20
QY 196 GAAGGCTGACGCGGAGATCTGAGAGACCAACCGGACATATACGATTTTCGAGCA 255
Db 21 GluGlyLeuThrArgGluIleLeuArgGluGlnProAspAsnIleProAlaPheAla 40
QY 256 GCCTATTTCGAGAGCCTTCTAGAGAAAGAGAGAAACCAACTTTGATCCAGAGATGG 315
Db 41 AlaTyrPheGluSerLeuLeuGluLysArgGluLysThrAsnPheAspProAlaGluT 60
QY 316 GGGAGTAAGTAGAAGCCGCTTCTATAACAATCATGCAATTCGAGGAGCAAGAACCACT 375
Db 61 GlySerLysValGluAspArgPheTyrAsnAsnHisAlaPheGluGlnGluProPro 80
QY 376 GAGAAAGTCACTCTTAACAAGAGCTCTCAGATATCTGGGAAGGAGGAGAGACATCA 435
Db 81 GluLysSerAspProLysGlnGluGluSerGlnIleSerGlyLysGluGluGluThrSer 100
QY 436 GTCAACATCTTAGACTCTCTGAGGAGAGATAAGGAGAAAGAGAGGTTGCTGTGTCAA 495
Db 101 ValThrIleLeuAspSerSerGluGluAspLysGluGluValAlaAlaValLys 120
QY 496 ATCCAGCTGCCTTCGGGGGACACATAGCCAGAGAGGAGGAGGAGGAGGAGGAGGAGG 555
Db 121 IleGlnAlaAlaPheArgGlyHisIleAlaArgGluGluAlaLysMetLysThrAsn 140
QY 556 AGTCTTCAAAATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 589
Db 141 SerLeuGlnAsnGluGluLysGluGluAsnLys 151

RESULT 4

US-09-513-999C-4914
; Sequence 4914, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59 US2, REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 4914
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 77
; OTHER INFORMATION: Xaa=His or Leu or Gln
US-09-513-999C-4914

Alignment Scores:
Pred. No.: 7,47e-64 Length: 119
Score: 76.00 Matches: 76
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.76% Indels: 0
DB: 4 Gaps: 0

AF334735 (1-954) x US-09-513-999C-4914 (1-119)

QY 136 ATGTCGATTCATCTCCACACCCACTACCGAATTCACAGATTTGGGAATCTTCTT 195
Db 1 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeu 20
QY 196 GAAGGCTGACGCGGAGATCTGAGAGACCAACCGGACATATACGATTTTCGAGCA 255
Db 21 GluGlyLeuThrArgGluIleLeuArgGluGlnProAspAsnIleProAlaPheAla 40
QY 256 GCCTATTTCGAGAGCCTTCTAGAGAAAGAGAGAAACCAACTTTGATCCAGAGATGG 315
Db 41 AlaTyrPheGluSerLeuLeuGluLysArgGluLysThrAsnPheAspProAlaGluT 60
QY 316 GGGAGTAAGTAGAAGCCGCTTCTATAACAATCATGCAATTCGAGGAG 363
Db 61 GlySerLysValGluAspArgPheTyrAsnAsnHisAlaPheGluGlu 76

RESULT 5

US-08-166-195A-50
; Sequence 50, Application US/08166195A
; Patent No. 5480799
; GENERAL INFORMATION:

;; APPLICANT: O'Rand, Michael G.
;; APPLICANT: Widgren, Esther E.
;; APPLICANT: Richardson, Richard T.
;; APPLICANT: Lea, Isabel
;; TITLE OF INVENTION: Sperm Antigen Corresponding to a
;; NUMBER OF SEQUENCES: 51
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Kenneth D. Sibley
;; STREET: P.O. Box 34009
;; CITY: Charlotte
;; STATE: No. 5480799th Carolina
;; COUNTRY: USA
;; ZIP: 28234
;; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/166,195A
FILING DATE: 10 DEC 1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470/73
TELEPHONE: 919-881-3140
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 145 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-166-195A-50

Alignment Scores:
Pred. No.: 2,3e-33 Length: 145
Score: 44.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.33% Indels: 0
DB: 1 Gaps: 0

AF334735 (1-954) x US-08-166-195A-50 (1-145)

QY 136 ATGTCGATTCATCTCCACACCCACTACCGAATTCACAGGATTTCGGAATCTCTT 195
Db 1 MetSerIleProPheSerAsnThrHisTyArgIleProGlnGlyPheGlyAsnLeu 20
QY 196 GAAGGCTGACGCGAGATTCTGAGAGACGACGACCAATATACCGCTTTTCGACGA 255
Db 21 GluGlyLeuThrArgGluLeuArgGluGlnProAspAsnIleProAlaPheAla 40

QY 256 GCCTATTTTGAG 267
Db 41 AlaTyPheGlu 44

RESULT 6
US-08-436-772-50
Sequence 50, Application US/08436772
Patent No. 5814456
GENERAL INFORMATION:
APPLICANT: O'Rand, Michael G.
APPLICANT: Widgren, Esther E.
APPLICANT: Richardson, Richard T.
APPLICANT: Lea, Isabel
TITLE OF INVENTION: Sperm Antigen Corresponding to a Sperm
TITLE OF INVENTION: Zona Binding Protein Autoantigenic Epitope
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenneth D. Sibley
STREET: P.O. Box 34009
CITY: Charlotte
STATE: No. 5814456th Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436,772
FILING DATE: 08-MAY-1995

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-73B
TELEPHONE: 919-881-3140
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 146 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-436-772-50
Alignment Scores:
Pred. No.: 2,29e-33 Length: 146
Score: 44.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.33% Indels: 0
DB: 2 Gaps: 0
AF334735 (1-954) x US-08-436-772-50 (1-146)
QY 136 ATGTCGATTCATCTCCACACCCACTACCGAATTCACAGGATTTCGGAATCTCTT 195
Db 1 MetSerIleProPheSerAsnThrHisTyArgIleProGlnGlyPheGlyAsnLeu 20
QY 196 GAAGGCTGACGCGAGATTCTGAGAGACGACGACCAATATACCGCTTTTCGACGA 255
Db 21 GluGlyLeuThrArgGluLeuArgGluGlnProAspAsnIleProAlaPheAla 40
QY 256 GCCTATTTTGAG 267
Db 41 AlaTyPheGlu 44
RESULT 7
US-08-436-883B-50
Sequence 50, Application US/08436883B
Patent No. 5820861
GENERAL INFORMATION:
APPLICANT: O'Rand, Michael G.
APPLICANT: Widgren, Esther E.
APPLICANT: Richardson, Richard T.
APPLICANT: Lea, Isabel
TITLE OF INVENTION: Sperm Antigen Corresponding to a Sperm
TITLE OF INVENTION: Zona Binding Protein Autoantigenic Epitope
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenneth D. Sibley
STREET: P.O. Box 34009
CITY: Charlotte
STATE: No. 5820861th Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436,883B
FILING DATE: 08-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-73C
TELEPHONE: 919-420-2200

```
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 146 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-436-883B-50
Alignment Scores:
Pred. No.: 2,298-33 Length: 146
Score: 44.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.33% Indels: 0
DB: 2 Gaps: 0
AF334735 (1-954) x US-08-436-883B-50 (1-146)
QY 136 ATGTCGATTCATTCCTCCACACCCACTACCGAATTCACAGGATTGGGAATCTTCTT 195
Db 1 MetSerIleProPheSerAsnThrHisTyArgIleProGlnGlyPheGlyAsnLeuLeu 20
QY 196 GAAGGGCTGACGCGAGATCTTGAGAGACCAACCGGACAAATATACAGCTTTTGCGCA 255
Db 21 GluGlyLeuThrArgGluIleLeuArgGluGlnProAspAsnIleProAlaPheAlaA1a 40
QY 256 GCTATTATTGAG 267
Db 41 AlaTyPheGlu 44
RESULT 8
US-08-166-195A-51
; Sequence 51, Application US/08166195A
; Patent No. 5480799
; GENERAL INFORMATION:
; APPLICANT: O'Rand, Michael G.
; APPLICANT: Widgren, Esther E.
; APPLICANT: Richardson, Richard T.
; APPLICANT: Lea, Isabel
; TITLE OF INVENTION: Sperm Antigen Corresponding to a
; TITLE OF INVENTION: Sperm Zona Binding Protein Autoantigenic Epitope
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenneth D. Sibley
; STREET: P.O. Box 34009
; CITY: Charlotte
; STATE: No. 5480799th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/166,195A
; FILING DATE: 10 DEC 1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470/73
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 148 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: linear
US-08-436-883B-50
Alignment Scores:
Pred. No.: 2,298-33 Length: 148
Score: 44.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
; MOLECULE TYPE: protein
US-08-166-195A-51
Alignment Scores:
Pred. No.: 2,298-33 Length: 148
Score: 44.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.33% Indels: 0
DB: 1 Gaps: 0
AF334735 (1-954) x US-08-166-195A-51 (1-148)
QY 136 ATGTCGATTCATTCCTCCACACCCACTACCGAATTCACAGGATTGGGAATCTTCTT 195
Db 1 MetSerIleProPheSerAsnThrHisTyArgIleProGlnGlyPheGlyAsnLeuLeu 20
QY 196 GAAGGGCTGACGCGAGATCTTGAGAGACCAACCGGACAAATATACAGCTTTTGCGCA 255
Db 21 GluGlyLeuThrArgGluIleLeuArgGluGlnProAspAsnIleProAlaPheAlaA1a 40
QY 256 GCTATTATTGAG 267
Db 41 AlaTyPheGlu 44
RESULT 9
US-08-436-772-51
; Sequence 51, Application US/08436772
; Patent No. 5814456
; GENERAL INFORMATION:
; APPLICANT: O'Rand, Michael G.
; APPLICANT: Widgren, Esther E.
; APPLICANT: Richardson, Richard T.
; APPLICANT: Lea, Isabel
; TITLE OF INVENTION: Sperm Antigen Corresponding to a Sperm
; TITLE OF INVENTION: Zona Binding Protein Autoantigenic Epitope
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenneth D. Sibley
; STREET: P.O. Box 34009
; CITY: Charlotte
; STATE: No. 5814456th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,772
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-73B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 148 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-436-772-51
Alignment Scores:
Pred. No.: 2,298-33 Length: 148
Score: 44.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.33% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x US-08-436-772-51 (1-148)

QY 136 ATGTCGATTCATTCTCCACACCCACTACCGAATTCACAGGATTTCGGAATCTTCTT 195
|||||
Db 1 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeu 20
|||||
QY 196 GAAGGGCTGACGCGAGATTCTGAGAGACCAACCGGACAATATACCACTTTTCAGCA 255
|||||
Db 21 GluGlyLeuThrArgGluIleLeuArgGluGlnProAspAsnIleProAlaPheAla 40
|||||
QY 256 GCCTATTTTGGAG 267
|||||
Db 41 AlatyPheGlu 44
|||||

RESULT 10

US-08-436-883B-51
; Sequence 51, Application US/08436883B
; Patent No. 5820861
; GENERAL INFORMATION:
; APPLICANT: O'Rand, Michael G.
; APPLICANT: Widgren, Esther E.
; APPLICANT: Richardson, Richard T.
; APPLICANT: Lea, Isabel
; TITLE OF INVENTION: Sperm Antigen Corresponding to a Sperm
; TITLE OF INVENTION: Zona Binding Protein Autoantigenic Epitope
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenneth D. Sibley
; STREET: P.O. Box 34009
; CITY: Charlotte
; STATE: No. 5820861th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,883B
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-73C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-420-2200
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 148 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-436-883B-51

Alignment Scores:
Pred. No.: 2,29e-33 Length: 148
Score: 44.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.33% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x US-08-436-883B-51 (1-148)

QY 136 ATGTCGATTCATTCTCCACACCCACTACCGAATTCACAGGATTTCGGAATCTTCTT 195

Db 1 MetSerIleProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeu 20
|||||
QY 196 GAAGGGCTGACGCGAGATTCTGAGAGACCAACCGGACAATATACCACTTTTCAGCA 255
|||||
Db 21 GluGlyLeuThrArgGluIleLeuArgGluGlnProAspAsnIleProAlaPheAla 40
|||||
QY 256 GCCTATTTTGGAG 267
|||||
Db 41 AlatyPheGlu 44
|||||

RESULT 11

US-08-436-772-57
; Sequence 57, Application US/08436772
; Patent No. 5814456
; GENERAL INFORMATION:
; APPLICANT: O'Rand, Michael G.
; APPLICANT: Widgren, Esther E.
; APPLICANT: Richardson, Richard T.
; APPLICANT: Lea, Isabel
; TITLE OF INVENTION: Sperm Antigen Corresponding to a Sperm
; TITLE OF INVENTION: Zona Binding Protein Autoantigenic Epitope
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenneth D. Sibley
; STREET: P.O. Box 34009
; CITY: Charlotte
; STATE: No. 5814456th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,772
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-73B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-436-772-57

Alignment Scores:
Pred. No.: 5.31e-18 Length: 28
Score: 28.00 Matches: 28
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.12% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x US-08-436-772-57 (1-28)

QY 298 TTTGATCCAGCAGAAATGGGGAGTAAGTAGAGACCGCTTCTATACCAATCATCATTC 357
|||||

Db 1 PheAspProAlaGluTyrGlySerIysValGluAspArgPheTyrAsnHisAlaPhe 20
|||||

QY 358 GAGGAGCAAGAACCCACCTGAGAAA 381
|||||

Db 21 GluGluGlnGluProProGluLys 28
|||||

```
RESULT 12
US-08-436-883B-57
; Sequence 57, Application US/08436883B
; Patent No. 5820861
; GENERAL INFORMATION:
; APPLICANT: O'Rand, Michael G.
; APPLICANT: Widgren, Esther E.
; APPLICANT: Richardson, Richard T.
; APPLICANT: Lea, Isabel
; TITLE OF INVENTION: Sperm Antigen Corresponding to a Sperm
; TITLE OF INVENTION: Zona Binding Protein Autoantigenic Epitope
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenneth D. Sibley
; STREET: P.O. Box 34009
; CITY: Charlotte
; STATE: No. 5820861th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,883B
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-73C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-420-2200
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-436-883B-57

Alignment Scores:
Pred. No.: 5,31e-18 Length: 28
Score: 28.00 Matches: 28
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.12% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x US-08-436-883B-57 (1-28)
QY 298 TTTGATCCAGCAATGGGAGTAAGTGAAGACCGCTTCATACATCATCATTC 357
Db 1 PheAspProAlaGluTrpGlySerLysValGluAspArgPheTyraAsnHisAlaPhe 20
QY 358 GAGGAGCAACACCACTTGAA 381
Db 21 GluGluGluGluProGluLys 28

RESULT 13
US-08-436-772-52
; Sequence 52, Application US/08436772
; Patent No. 5814456
; GENERAL INFORMATION:
; APPLICANT: O'Rand, Michael G.
; APPLICANT: Widgren, Esther E.
; APPLICANT: Richardson, Richard T.
; APPLICANT: Lea, Isabel
; TITLE OF INVENTION: Sperm Antigen Corresponding to a Sperm
; TITLE OF INVENTION: Zona Binding Protein Autoantigenic Epitope
```

```
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenneth D. Sibley
; STREET: P.O. Box 34009
; CITY: Charlotte
; STATE: No. 5814456th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,772
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-73B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-436-772-52

Alignment Scores:
Pred. No.: 3,91e-15 Length: 25
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.14% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x US-08-436-772-52 (1-25)
QY 145 CCATTCTCCACACCCACTACCGAATTCACAGGATTTGGGAATCTTCTTGAAGGCTG 204
Db 1 ProPheSerAsnThrHisTyraGileProGlnGlyPheGlyAsnLeuGluGlyLeu 20
QY 205 ACACGGAGATTCTG 219
Db 21 ThrArgGluIleLeu 25

RESULT 14
US-08-436-883B-52
; Sequence 52, Application US/08436883B
; Patent No. 5820861
; GENERAL INFORMATION:
; APPLICANT: O'Rand, Michael G.
; APPLICANT: Widgren, Esther E.
; APPLICANT: Richardson, Richard T.
; APPLICANT: Lea, Isabel
; TITLE OF INVENTION: Sperm Antigen Corresponding to a Sperm
; TITLE OF INVENTION: Zona Binding Protein Autoantigenic Epitope
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenneth D. Sibley
; STREET: P.O. Box 34009
; CITY: Charlotte
; STATE: No. 5820861th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436,883B
FILING DATE: 08-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-73C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-436-883B-52

Alignment Scores:
Pred. No.: 3,91e-15 Length: 25
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.14% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x US-08-436-883B-52 (1-25)

QY 145 CCATCTCCACACCCACTACCAATCCACAGGATTTGGGATCTCTTGAAGGGCTG 204
Db 1 ProPheSerAsnThrHisTyrArgIleProGlnGlyPheGlyAsnLeuLeuGluGlyLeu 20
QY 205 ACACGCGAGATTCGTG 219
Db 21 ThrArgGluIleLeu 25

RESULT 15

US-08-436-772-60
Sequence 60, Application US/08436772
Patent No. 5814456
GENERAL INFORMATION:
APPLICANT: O'Rand, Michael G.
APPLICANT: Widgren, Esther E.
APPLICANT: Richardson, Richard T.
APPLICANT: Lea, Isabel
TITLE OF INVENTION: Sperm Antigen Corresponding to a Sperm
TITLE OF INVENTION: Zona Binding Protein Autoantigenic Epitope
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenneth D. Sibley
STREET: P.O. Box 34009
CITY: Charlotte
STATE: No. 5814456th Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436,772
FILING DATE: 08-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-73B
TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-881-3140
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-436-772-60
Alignment Scores:
Pred. No.: 2,61e-11 Length: 21
Score: 21.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.84% Indels: 0
DB: 2 Gaps: 0

AF334735 (1-954) x US-08-436-772-60 (1-21)

QY 481 GTTGCTGCTGTCAAAATCCAACTCCCTTCGGGACACATAGCCAGAGGAGGCAAG 540
Db 1 ValAlaAlaValIleGlnAlaAlaPheArgGlyHisIleAlaArgGluGluAlaLys 20
QY 541 AAA 543
Db 21 Lys 21

Search completed: February 15, 2005, 02:31:47
Job time : 39.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 15, 2005, 02:28:57 ; Search time 147 Seconds
(without alignments)
4241.070 Million cell updates/sec

Title: AF334735
Perfect score: 307
Sequence: 1 TCGCCCTTCGCGCGCG.....AAAAAAAAAAAAAAAAAAAA 954

Scoring table:
Xgapop 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1376875 seqs, 326749119 residues
Word size: 1

Total number of hits satisfying chosen parameters: 2645598
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-Q=/cgn2_1/USFTO.spool_p/af334735/runat_14022005_160241_19433/app_query.fasta_1.1095
-DB=Published Applications AA -QMT=fastan -SUFFIX=Oligo.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USRR=AF334735 @CNG 1.1 199 @runat_14022005_160241_19433
-NCFU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pcp.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pcp.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description
1	11	3.6	157	15	US-10-424-599-166824	Sequence 166824,
2	11	3.6	236	15	US-10-424-599-162808	Sequence 162808,
3	10	3.3	113	15	US-10-437-963-121776	Sequence 121776,
4	9	2.9	63	15	US-10-424-599-211183	Sequence 211183,
5	9	2.9	104	15	US-10-424-599-174377	Sequence 174377,
6	9	2.9	105	16	US-10-437-963-162142	Sequence 162142,
7	9	2.9	126	15	US-10-424-599-247560	Sequence 247560,
8	9	2.9	137	15	US-10-424-599-249014	Sequence 249014,
9	9	2.9	202	15	US-10-424-599-249011	Sequence 249011,
10	9	2.9	592	14	US-10-100-294A-36	Sequence 36, Appl
11	9	2.9	602	15	US-10-104-047-3092	Sequence 3092, Ap
12	9	2.9	860	15	US-10-341-434-47	Sequence 47, Appl
13	9	2.9	860	15	US-10-416-592-2	Sequence 2, Appl
14	9	2.9	6620	15	US-10-080-334-290	Sequence 290, App
15	9	2.9	6620	16	US-10-408-765A-2291	Sequence 2291, Ap
16	9	2.9	7968	13	US-10-077-130-5	Sequence 5, Appl
17	9	2.6	23	10	US-09-986-480-311	Sequence 311, App
18	8	2.6	26	9	US-09-764-847-702	Sequence 702, App
19	8	2.6	26	14	US-10-092-154-702	Sequence 702, App
20	8	2.6	28	15	US-10-424-599-184595	Sequence 184595,
21	8	2.6	34	15	US-10-424-599-277778	Sequence 277778,
22	8	2.6	34	16	US-10-437-963-146931	Sequence 146931,
23	8	2.6	37	9	US-09-764-869-777	Sequence 777, App
24	8	2.6	37	14	US-10-091-504-777	Sequence 777, App
25	8	2.6	37	15	US-10-227-577-777	Sequence 777, App
26	8	2.6	38	16	US-10-437-963-105267	Sequence 105267,
27	8	2.6	39	15	US-10-424-599-208544	Sequence 208544,
28	8	2.6	39	15	US-10-424-599-247879	Sequence 247879,
29	8	2.6	40	16	US-10-437-963-109774	Sequence 109774,
30	8	2.6	42	16	US-10-437-963-188679	Sequence 188679,
31	8	2.6	43	16	US-10-437-963-156489	Sequence 156489,
32	8	2.6	44	15	US-10-424-599-195083	Sequence 195083,
33	8	2.6	44	16	US-10-291-226-91	Sequence 91, Appl
34	8	2.6	44	16	US-10-437-963-131469	Sequence 131469,
35	8	2.6	45	16	US-10-437-963-201534	Sequence 201534,
36	8	2.6	46	15	US-10-424-599-248344	Sequence 248344,
37	8	2.6	46	16	US-10-291-226-96	Sequence 96, Appl
38	8	2.6	46	16	US-10-437-963-154807	Sequence 154807,
39	8	2.6	47	9	US-09-814-122-48	Sequence 48, Appl
40	8	2.6	47	15	US-10-424-599-177226	Sequence 177226,
41	8	2.6	47	15	US-10-649-857-48	Sequence 48, Appl
42	8	2.6	48	15	US-10-424-599-170716	Sequence 170716,
43	8	2.6	48	15	US-10-424-599-254354	Sequence 254354,
44	8	2.6	48	15	US-10-424-599-256282	Sequence 256282,
45	8	2.6	49	15	US-10-424-599-254498	Sequence 254498,

ALIGNMENTS

RESULT 1
US-10-424-599-166824
; Sequence 166824, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 166824
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_121657C.1.pap
US-10-424-599-166824

Alignment Scores:
Pred. No.: 0.277 Length: 157
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.58% Indels: 0
DB: 15 Gaps: 0

AF334735 (1-954) x US-10-424-599-166824 (1-157)

QY 484 GCTGCTGTCAAATCCAAAGCTGCTTCGGGGA 516
Db 14 AlaAlaVallyslleGlnAlaAlaPheArgGly 24

RESULT 2

US-10-424-599-162808
; Sequence 162808, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 162808
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_118033C.1.pap
US-10-424-599-162808

Alignment Scores:
Pred. No.: 0.259 Length: 236
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.58% Indels: 0
DB: 15 Gaps: 0

AF334735 (1-954) x US-10-424-599-162808 (1-236)

QY 484 GCTGCTGTCAAATCCAAAGCTGCTTCGGGGA 516
Db 64 AlaAlaVallyslleGlnAlaAlaPheArgGly 74

RESULT 3

US-10-437-963-121776
; Sequence 121776, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazur, Brad
; APPLICANT: Li, Ping
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 121776
; LENGTH: 113

; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_24768C.1.pap
US-10-437-963-121776

Alignment Scores:
Pred. No.: 2.67 Length: 113
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.26% Indels: 0
DB: 16 Gaps: 0

AF334735 (1-954) x US-10-437-963-121776 (1-113)

QY 24 TTTTITTTTTTTTAAAGAAAAACGGTTA 53
Db 23 PhePhePheLeuLysLysLysArgLeu 32

RESULT 4

US-10-424-599-211183
; Sequence 211183, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 211183
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_32724C.1.pap
US-10-424-599-211183

Alignment Scores:
Pred. No.: 26.9 Length: 63
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.93% Indels: 0
DB: 15 Gaps: 0

AF334735 (1-954) x US-10-424-599-211183 (1-63)

QY 24 TTTTITTTTTTTTAAAGAAAAACGG 50
Db 20 PhePhePheLeuLysLysLysArg 28

RESULT 5

US-10-424-599-174377
; Sequence 174377, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 174377

FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 247560
LENGTH: 126
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_128481C.1.pap
US-10-424-599-174377

Alignment Scores:
Pred. No.: 24.7 Length: 104
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.93% Indels: 0
DB: 15 Gaps: 0

AF334735 (1-954) x US-10-424-599-174377 (1-104)

QY 27 TTTTITTTTTTAAAGAAAAACGGTTA 53
DB 10 PhePheLeuLysLysLysLysLysLysLys 18

RESULT 6

US-10-437-963-162142

Sequence 162142, Application US/10437963
Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 162142

LENGTH: 105

TYPE: PRT

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT4530_61262C.1.pap

US-10-437-963-162142

Alignment Scores:
Pred. No.: 24.7 Length: 105
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.93% Indels: 0
DB: 16 Gaps: 0

AF334735 (1-954) x US-10-437-963-162142 (1-105)

QY 928 ATCCCCCAAAAAAAAAAAAAAAAAA 954
DB 87 IleProGlnLysLysLysLysLysLys 95

RESULT 7

US-10-424-599-247560

Sequence 247560, Application US/10424599

Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 247560
LENGTH: 126
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_65576C.1.pap
US-10-424-599-247560

Alignment Scores:
Pred. No.: 23.9 Length: 126
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 15 Gaps: 0

AF334735 (1-954) x US-10-424-599-247560 (1-126)

QY 329 TCTACTTACTCCCCCATTCGTGGA 303

DB 16 SerThrLeuLeuProHisSerAlaGly 24

RESULT 8

US-10-424-599-249014

Sequence 249014, Application US/10424599

Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 249014

LENGTH: 137

TYPE: PRT

ORGANISM: Glycine max

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(137)

OTHER INFORMATION: unsure at all Xaa locations

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT3847_66890C.1.pap

US-10-424-599-249014

Alignment Scores:
Pred. No.: 23.6 Length: 137
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.93% Indels: 0
DB: 15 Gaps: 0

AF334735 (1-954) x US-10-424-599-249014 (1-137)

QY 487 GCTGTCAAAATCCAAAGCTTCGGG 513

DB 9 AlaValLysIleGlnAlaAlaPheArg 17

RESULT 9

US-10-424-599-249011

Sequence 249011, Application US/10424599

Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 249011
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_66888C.1.pap
US-10-424-599-249011

Alignment Scores:
Pred. No.: 22.1 Length: 202
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.93% Indels: 0
DB: 15 Gaps: 0

AF334735 (1-954) x US-10-424-599-249011 (1-202)

QY 487 GCTGTCAAATCCAGCTGCTTCGG 513
|||||
Db 77 AlavallyaileGlnAlaAlaPheArg 85

RESULT 10

US-10-100-294A-36
; Sequence 36, Application US/10100294A
; Publication No. US20030135877A1
; GENERAL INFORMATION:
; APPLICANT: HAERTEL, HEIKO A.
; APPLICANT: MITTENDORF, VOLKER
; APPLICANT: CHEN, RUOYING
; APPLICANT: SHANK, KARIN J.
; APPLICANT: RESKI, RALF
; TITLE OF INVENTION: SUGAR AND LIPID METABOLISM REGULATORS IN PLANTS
; FILE REFERENCE: 16313-0095
; CURRENT APPLICATION NUMBER: US/10/100,294A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/276,993
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 36
; LENGTH: 592
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-10-100-294A-36

Alignment Scores:
Pred. No.: 18.5 Length: 592
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.93% Indels: 0
DB: 14 Gaps: 0

AF334735 (1-954) x US-10-100-294A-36 (1-592)

QY 926 ACATCCCCCAAAAAAAAAAAAAAAAAA 952
|||||
Db 498 ThrSerProLysLysLysLysLys 506

RESULT 11

US-10-104-047-3092
; Sequence 3092, Application US/10104047
; Publication No. US20030236392A1

; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3092
; LENGTH: 602
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3092

Alignment Scores:
Pred. No.: 18.4 Length: 602
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.93% Indels: 0
DB: 15 Gaps: 0

AF334735 (1-954) x US-10-104-047-3092 (1-602)

QY 926 ACATCCCCCAAAAAAAAAAAAAAAAAA 952
|||||
Db 148 ThrSerProLysLysLysLysLys 156

RESULT 12

US-10-341-434-47
; Sequence 47, Application US/10341434
; Publication No. US20030215835A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
; FILE REFERENCE: 9U 204 205 R1
; CURRENT APPLICATION NUMBER: US/10/341,434
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/348,164
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/348,119
; PRIOR FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 47
; LENGTH: 860
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-341-434-47

Alignment Scores:
Pred. No.: 17.4 Length: 860
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.93% Indels: 0
DB: 15 Gaps: 0

AF334735 (1-954) x US-10-341-434-47 (1-860)

QY 926 ACATCCCCCAAAAAAAAAAAAAAAAAA 952
|||||
Db 148 ThrSerProLysLysLysLysLys 156

RESULT 13

US-10-416-592-2
; Sequence 2, Application US/10416592
; Publication No. US20040053291A1
; GENERAL INFORMATION:
; APPLICANT: TANG, Y. Tom
; APPLICANT: WARREN, Bridget A.

APPLICANT: HAFALIA, April J.A.
APPLICANT: ARVIZO, Chandra
APPLICANT: LU, Dyung Aina M.
APPLICANT: ELLIOTT, Vicki S.
APPLICANT: XU, Yuming
APPLICANT: LAL, Preeti G.
APPLICANT: YUE, Henry
TITLE OF INVENTION: NUCLEIC ACID-ASSOCIATED PROTEINS
FILE REFERENCE: PF-0847 USN
CURRENT APPLICATION NUMBER: US/10/416,592
CURRENT FILING DATE: 2003-05-12
PRIOR APPLICATION NUMBER: PCT/US01/43606
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/248,404
PRIOR FILING DATE: 2000-11-13
PRIOR APPLICATION NUMBER: US 60/249,825
PRIOR FILING DATE: 2000-11-16
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PERL Program
SEQ ID NO 2
LENGTH: 860
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040053291A1 1933163CD1
US-10-416-592-2

Alignment Scores:
Pred. No.: 17.4 Length: 860
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.93% Indels: 0
DB: 15 Gaps: 0

AF334735 (1-954) x US-10-416-592-2 (1-860)
QY 926 ACATCCCCCAAAAAAAAAAAAAAAAAAAAA 952
Db 148 ThrSerProlyslslyslslyslslysls 156

RESULT 14
US-10-080-334-290
Sequence 290, Application US/10080334
Publication No. US2004002584A1
GENERAL INFORMATION:
APPLICANT: Pena, Carol E. A.
APPLICANT: Shimkets, Richard A
APPLICANT: Li, Li
APPLICANT: Shenoy, Suresh G
APPLICANT: Kekuda, Ramesh
APPLICANT: Spytak, Kimberly A
APPLICANT: Vernet, Corine A. M.
APPLICANT: Malyankar, Uriel M
APPLICANT: Guo, Xiaojia
APPLICANT: Gusev, Vladimir Y
APPLICANT: Casman, Stacie J
APPLICANT: Boldog, Ferenc L
APPLICANT: Furtak, Katarzyna
APPLICANT: Tchernev, Velizar T
APPLICANT: Gangolli, Baha A
APPLICANT: Padigaru, Muralidhara
APPLICANT: Liu, Xisohong
APPLICANT: Baumgartner, Jason C.
APPLICANT: Gerlach, Valerie
APPLICANT: Spaderna, Steven K
APPLICANT: Zerhusen, Bryan D
TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
FILE REFERENCE: 21402-275
CURRENT APPLICATION NUMBER: US/10/080,334

CURRENT FILING DATE: 2002-02-21
PRIOR APPLICATION NUMBER: 60/270,523
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/322,712
PRIOR FILING DATE: 2001-09-17
PRIOR APPLICATION NUMBER: 60/311,980
PRIOR FILING DATE: 2001-08-13
PRIOR APPLICATION NUMBER: 60/330,307
PRIOR FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: 60/278,796
PRIOR FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: 60/281,521
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/276,677
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/311,595
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/270,220
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/274,295
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/318,526
PRIOR FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/286,548
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: 60/291,765
PRIOR FILING DATE: 2001-05-17
PRIOR APPLICATION NUMBER: 60/270,797
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 60/276,400
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/270,810
PRIOR FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 388
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 290
LENGTH: 6620
TYPE: PRT
ORGANISM: Homo sapiens
US-10-080-334-290

Alignment Scores:
Pred. No.: 12.3 Length: 6620
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.93% Indels: 0
DB: 15 Gaps: 0

AF334735 (1-954) x US-10-080-334-290 (1-6620)
QY 484 GCTGCTGCAAAATCCAGCTGCCTTC 510
Db 4875 AlaAlaVallyslleGlnAlaAlaPhe 4883

RESULT 15
US-10-408-765A-2291
Sequence 2291, Application US/10408765A
Publication No. US20040101874A1
GENERAL INFORMATION:
APPLICANT: Ghosh, Soumitra S.
APPLICANT: Fahy, Eoin D.
APPLICANT: Zhang, Bing
APPLICANT: Gibson, Bradford W.
APPLICANT: Taylor, Steven W.
APPLICANT: Glenn, Gary M.
APPLICANT: Warnock, Dale E.
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
FILE REFERENCE: 660088.465
CURRENT APPLICATION NUMBER: US/10/408,765A
CURRENT FILING DATE: 2003-04-04
NUMBER OF SEQ ID NOS: 3077

```

; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2291
; LENGTH: 6620
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2291

Alignment Scores:
Pred. No.:      12.3      Length:      6620
Score:          9.00      Matches:      9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:      2.93% Indels:      0
DB:              16      Gaps:      0

AF334735 (1-954) x US-10-408-765A-2291 (1-6620)

QY 484 GCTGCTGCAAAATCCAAGTCCTTC 510
    |||||
Db 4875 AlaAlaValLysIleGlnAlaAlaPhe 4883

```

Search completed: February 15, 2005, 02:51:50
Job time : 160 secs